View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Sound Ideas

University of Puget Sound
Sound Ideas

School of Occupational Master's Capstone Projects | Occupational Therapy, School of

5-2016

Identifying and Implementing Assessments of Upper Extremity Motor Control for Patients with Stroke or Parkinson's Disease

Julia Albright
University of Puget Sound

Kayla Karelsen
University of Puget Sound

Allison Lucas
University of Puget Sound

Follow this and additional works at: http://soundideas.pugetsound.edu/ot_capstone


Part of the Occupational Therapy Commons

Recommended Citation

Albright, Julia; Karelsen, Kayla; and Lucas, Allison, "Identifying and Implementing Assessments of Upper Extremity Motor Control for Patients with Stroke or Parkinson's Disease" (2016). *School of Occupational Master's Capstone Projects*. 2.
http://soundideas.pugetsound.edu/ot_capstone/2

This Article is brought to you for free and open access by the Occupational Therapy, School of at Sound Ideas. It has been accepted for inclusion in School of Occupational Master's Capstone Projects by an authorized administrator of Sound Ideas. For more information, please contact soundideas@pugetsound.edu.

Identifying and Implementing Assessments of Upper Extremity Motor Control for Patients with Stroke or Parkinson's Disease

May 2016

This evidence project, submitted by

Julia Albright, OTS
Kayla Karelsen, OTS
Allison Lucas, OTS

has been approved and accepted
in partial fulfillment of the requirements for the degree of
Master of Science in Occupational Therapy from the University of Puget Sound.

_____
Project Chairperson: Wendell Nakamura, MOT, OTR/L

_____ _____
OT635/636 Instructors: George Tomlin, PhD, OTR/L, FAOTA; Sue Doyle, PhD, OTR/L, CFE

_____
Director, Occupational Therapy Program: Yvonne Swinth, PhD, OTR/L, FAOTA

_____
Dean of Graduate Studies: Sunil Kukreja, PhD

Keywords: stroke, Parkinson's disease, assessments, upper extremity, knowledge translation, occupational therapy

Abstract

The authors collaborated with a team of clinicians at a skilled nursing facility (SNF) in Gig Harbor, WA to identify the need for measures of UE functional use that are most appropriate to utilize with the two largest client populations at their facility, stroke and Parkinson's disease (PD). In order to meet the clinical utility characteristics identified as important by the clinicians, the authors searched two databases that contain overviews of assessments. This search identified five stroke assessments and two PD assessments that met the clinical utility parameters outlined by the clinicians. The authors then conducted a systematic literature search to identify any relevant studies pertaining to each assessment. These studies were reviewed and the results presented in the form of a critical appraisal of the topic (CAT) that included the purpose, measure, population, psychometric characteristics, results and limitations of the studies. The initial search for articles yielded 869 articles, 33 of which met the inclusion criteria. The authors recommend use of the Chedoke Arm and Hand Activity Inventory-9 assessment with clients post-stroke due to its robust evidence and wide variety of levels of impairments post-stroke included in the studies, and use of the Muscle Disorder Society's Unified Parkinson's Disease Rating Scale assessment with clients with PD due to its inclusion of functional tasks and higher volume of research.

The knowledge translation implementation phase of the project involved an in-service (interactive demonstration and scoring) and two brochures outlining the findings related to each assessment. A follow-up survey measured the effectiveness of the in-service and the value of the research to clinicians' occupational therapy practice at their clinic. The survey results indicated that the clinicians found the research helpful and the knowledge translation process efficient. In addition, they indicated that it is likely that they will implement the CAHAI and the MDS-UPDRS assessments at their facility.

Executive Summary

Occupational therapy clinicians at a skilled nursing rehabilitation clinic collaborated with Master of Science in Occupational Therapy students from the University of Puget Sound on a research project. The aim of the project was to answer a clinical question and streamline the knowledge translation process. The project was completed over a period of approximately 9 months and it was broken down into the following two parts: 1) a systematic review of the literature to recommend specific assessments, and 2) assisting the clinicians with implementing the recommendations. The researchers began by interviewing the clinicians to learn about their needs and interests. Their question was: Which assessment is the most accurate and sensitive outcome measure of voluntary motor control in the upper extremity (UE) that can be used in their setting? The question was multifaceted and included applicability to the primary diagnoses seen in the facility, psychometric data properties to support its use in this setting, focus on functional use of the UE versus impairment, and feasibility and practicality to administer widely within the preexisting framework in the facility.

The outcome measure used by occupational therapists at the facility was the Modified Barthel Index. The occupational therapists were dissatisfied as they felt that this measure did not identify changes in UE function, an area that they spent significant time addressing in treatment. To be clinically useful, the clinicians identified the clinical utility parameters of the measurement tool (e.g., under $300 and <30 minutes to administer and score) as being important factors to consider in the selection of the assessment, as well as its suitability for persons with stroke or Parkinson's disease, the two largest client populations seen by the facility. The systematic literature review was completed in two stages: identifying assessments for consideration that met the clinical utility parameters, and investigating the psychometric properties of the identified assessments. Two databases of assessments were searched to find measures of UE function, which were then narrowed down by the clinical utility factors outlined by the clinicians. In total, five assessments for patients with stroke and two assessments for patients with Parkinson's disease were included in the search for articles. Thirty-three articles from ten databases were included in the critical appraisal of the topic (CAT) to allow the authors to analyze and compare the psychometric

properties of each assessment. Some limitations that negatively affected the generalizability of the studies' findings included excluding participants based on their cognitive impairment, level of motor return, and severity of condition symptoms.

The researchers concluded that the Chedoke Arm and Hand Activity Inventory (CAHAI-9) and the Movement Disorder Society's Unified Parkinson's Disease Rating Scale (MDS-UPDRS) are both measures that have strong psychometric properties including demonstrated high validity, reliability, and sensitivity for measuring upper extremity functional use in patients with mild to severe impairments as a result of stroke or PD, respectively. The two assessments also meet the specific needs of the clinicians in terms of cost, time to administer, and ability to measure functional use of the upper extremity.

The next phase of the project was to identify methods to support the implementation of the findings of the CAT (i.e., utilization of the two assessments) by the collaborating clinicians. An in-service was provided to the clinicians with the aims of familiarizing them with the assessments and providing them with an opportunity to have their concerns addressed. The in-service included a demonstration of how to administer the CAHAI-9 and an opportunity to practice scoring it. The authors left the clinicians with a CAHAI-9 manual, a cost estimate for creating a CAHAI-9 test kit, and two brochures that each included an overview of the CAHAI-9 and the MDS-UPDRS. During the in-service, the clinicians identified possible concerns, such as how cognitive impairments would affect performance, and how the assessment might be used with Medicare G-codes. At the conclusion of the in-service, the rehabilitation director told the other clinicians that she would like to set up a meeting with them soon to discuss the CAHAI-9 further and take the next steps necessary to implement the use of the assessment in their clinic.

A follow-up survey was used to further measure the outcome of the knowledge translation process. The clinicians indicated on the survey that it is likely that they will implement the CAHAI and the MDS-UPDRS assessments at their facility. When asked, they gave no suggestions for ways to improve the process of knowledge translation. They indicated that the research was useful to them and the knowledge translation process was efficient.

FOCUS QUESTION

Which of the selected assessments is the most accurate and sensitive outcome measure of functional abilities of the upper extremity for a skilled nursing-based rehabilitation setting? This question has multiple other facets included under it: is this tool applicable to the primary diagnoses seen in the facility (stroke and Parkinson's disease), does the measure have psychometric properties to support its use in this setting, and is it feasible and practical to administer widely within the preexisting framework in the facility?

CLINICAL SCENARIO

A director of rehabilitation and her team at a skilled nursing facility rehabilitation clinic is wondering which upper extremity functional use/ability measure for clients with neurological upper extremity impairment is the most accurate and sensitive. Approximately 35% of the clients treated at the facility have been diagnosed as either post-stroke (15%) or as having Parkinson's disease (20%). It is important to note that the clinicians only treat PD secondary to other diagnoses such as falls. The clinicians want a tool that is sensitive enough to measure changes in upper extremity voluntary motor control during functional use for clients who are post-stroke and for clients with Parkinson's disease, both at intake and throughout the intervention process. They are displeased with their current outcome measure, the Modified Barthel Index, because it does not identify changes in upper extremity performance and does not meet their needs. The clinicians identified that the clinical utility of the assessment (e.g., under $300 and <30 minutes to administer and score, and ability to measure motor control in the context of function) as an important factor to consider in the selection of the assessment. Because such a large portion of the client population seen at the facility has a diagnosis of either stroke or PD, the authors analyzed and summarized research on both of those populations in this critical appraisal of the topic (CAT).

REVIEW PROCESS

Rationale for Stage 1, Identifying Assessments for Consideration

The authors first narrowed the search by assessment type to provide a more focused understanding of assessments pertinent to upper extremity motor control in patients who are post-stroke or have PD. The authors conducted searches of the two diagnoses separately.

Search Strategy for Stroke Assessments. A number of assessments measure upper extremity ability or functional use in individuals with stroke. However, it was important to offer the clinicians a manageable sample of assessments to compare and from which they might select the one that best suits their needs. Stroke Engine and Rehab Measures are two trusted databases designed to bridge the gap between research and clinical practice (Korner-Bitensky & Rochette, n.d.; Rehabilitation Measures Database, n.d.). Health professionals often use these sites to choose appropriate assessments and interventions, many of which have been researched specifically with individuals who are post-stroke. Additionally, the two databases explicitly list clinical utility features so assessments that met the needs of the clinicians could be identified prior to searching for studies that evaluate the psychometric features of the measures.

An initial search of the Rehab Measures database produced 27 potential assessments for inclusion and a search of the Stroke Engine database produced 13. There were 8 assessments that were common to both databases. As a result, there were a total of 32 distinct assessments of motor control for people post-stroke that were identified for consideration. Assessments were further narrowed down by applying specific exclusion criteria (based on the clinicians' stated needs for clinical utility and the information available related to the psychometric characteristics of the measure) to ensure that each assessment included in the literature review (Stage 2) would meet the clinicians' needs.

Exclusion Criteria:

- < 2 articles on psychometric data of assessment
- Insufficient information available (e.g., no cost information available)
- > 30 minutes to administer and score

- > $300
- Assessments that measure client factor level impairments only
- Client-reported outcome measures
- > 50% of task items included gait/mobility
- Assessments appropriate for a pediatric population only

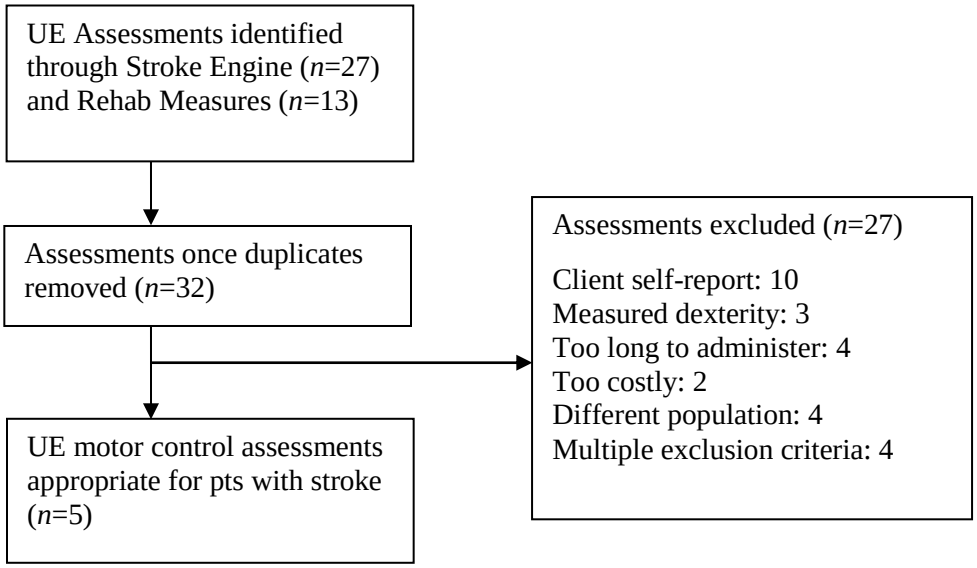

Figure 1: Flow chart of the process of identifying assessments for stroke to include in this critical appraisal.

Search Strategy for PD assessments. There are a limited number of assessments that measure upper extremity motor control in individuals with PD identified by our initial searches. A database comparable to Stroke Engine that lists information about assessments for individuals with PD was not found. A search of Rehab Measures yielded eight assessments appropriate for use with individuals who have PD. The authors expanded the search for assessments that are appropriate for use with individuals with PD by searching for systematic reviews (see Stage 2). In addition, 7 other assessments that focus on upper extremity motor control were selected from a recent systematic review (Proud et al., 2015). The systematic review was found in PubMed using the following search string: "upper extremity" AND "Parkinson" AND "systematic review". The systematic review by Proud et al. (2015) was selected because it was published within one year of writing this CAT and its authors compared assessments that

measure upper extremity functioning in patients with PD. Using the same strategy as used for stroke, PD assessments were also narrowed down by applying specific exclusion criteria to ensure that each assessment included in the article search (Part 2) would meet the clinicians' needs.

Exclusion Criteria:

- < 2 articles on psychometric data of assessment
- Insufficient information available (e.g., no cost information available)
- Replaced by a psychometrically superior version
- > 30 minutes to administer and score
- > $300
- Assessments that measure dyskinesia or tremor only
- Client-reported outcome measures
- Primary utility is in pharmaceutical research
- Assessments appropriate for a pediatric population only

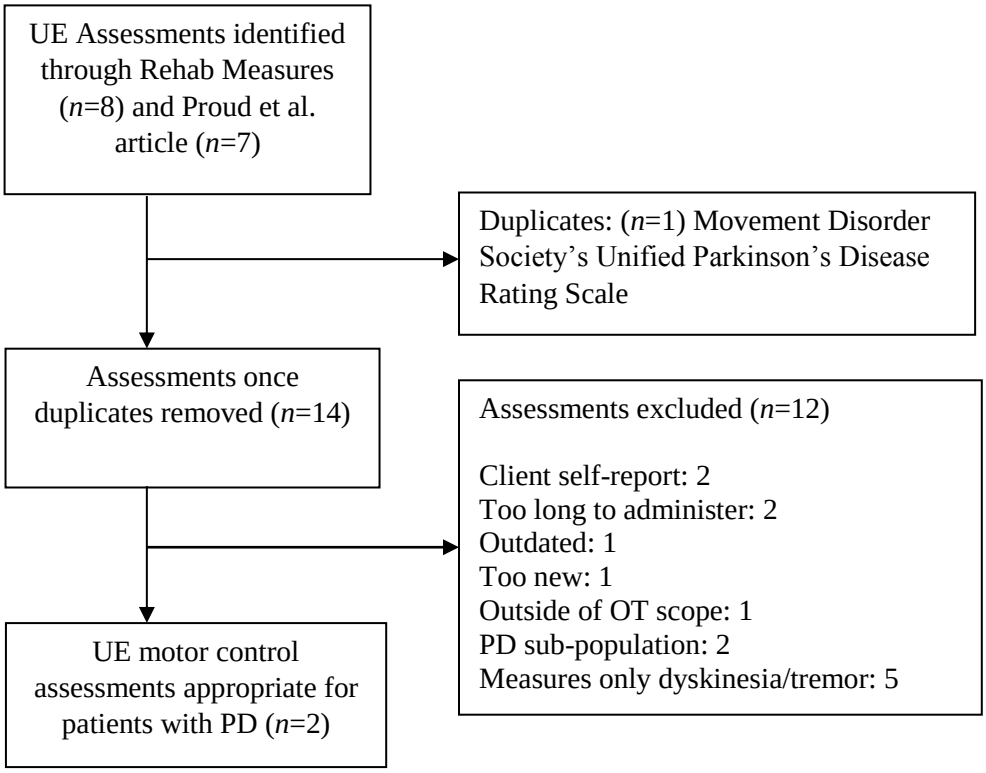

Figure 2: Flow chart of the process of identifying assessments for PD to include in this critical appraisal.

Results of Stage 1 search strategy:

One assessment from Stroke Engine and four assessments from Rehab Measures that measure upper extremity motor control in patients post-stroke met the inclusion and exclusion criteria.

The following five assessments that measure upper extremity motor control in patients who are post-stroke met the inclusion and exclusion criteria:

1. Action Research Arm Test (ARAT)
2. Streamlined Wolf Motor Function Test (S-WMFT)
3. The Chedoke Arm and Hand Activity Inventory (CAHAI)
4. Arm Motor Abilities Test (AMAT)
5. Motor Evaluation Scale for Upper Extremity in Stroke Patients (MESUPES)

Only one assessment from the Proud et al. (2015) article met the search criteria and was deemed appropriate for the collaborating clinicians to administer to patients with PD. Searching Rehab Measures for assessments for patients with PD that met all of the inclusion and exclusion criteria uncovered only one assessment. The following two assessments that measure upper extremity motor control in patients with PD met the inclusion and exclusion criteria:

1. Rush Dyskinesia Rating Scale (RDRS)
2. Movement Disorder Society's Unified Parkinson's Disease Rating Scale (MDS-UPDRS)

Following Stage 1, the assessment search, the first research question was answered. Seven assessments that were applicable to the primary diagnoses seen in the facility and met the clinical utility needs identified by the clinicians were identified.

Rationale for Stage 2, Investigating the Psychometric Properties of Select Assessments

Occupational therapists and other health care professionals are expected to provide care guided by best practices to which knowledge translation is closely tied. This critical appraisal of the topic was initiated in an attempt to facilitate knowledge translation. The process of knowledge translation attempts to eliminate the underutilization of evidence-based research to support systems of healthcare (Center on Knowledge Translation for Disability and Rehabilitation Research [KTDRR], 2005). Knowledge

translation involves ethically applying research to change and improve the quality and effectiveness of care that patients receive. The selection of an assessment based on diagnosis and clinical utility alone is insufficient to implement an assessment into practice. It is important to consider how different assessments may be more or less appropriate for particular client populations. For example, the sensitivity of some assessments may differ depending on the severity of a person's health complications. This CAT provided a more in-depth analysis of the psychometric properties of each assessment as they were administered to specific populations.

Search Strategy for Stroke Articles. Databases were chosen based on their relevance to the field of occupational therapy with the expectation that they would return many results. Key terms were chosen based on the assessments included and what information was relevant to the needs of the collaborating clinicians (see Table 1). When necessary, filters were added to the search in order to narrow down results to garner more relevant articles that fit the inclusion criteria (See Table 3).

Stroke Related Articles Search Process

Table 1

Search Terms for Stroke-Related Studies

Key Terms	**Synonyms**	**Alternate Spellings**
Action Research Arm Test		ARAT
Arm Motor Abilities Test	Arm Mobility Arm Test	AMAT
The Chedoke Arm and Hand Activity Inventory		CAHAI, Chedoke
Motor Evaluation Scale for Upper Extremity in Stroke Patients		MESUPES
Psychometric	Clinimetric, reliable, valid, sensitivity, specificity	reliab*, valid*
Streamlined Wolf Motor Function Test		S-WMFT, SWMFT
Upper Extremity	Arm, hand, upper limb,	UE

Table 2

Databases Searched for Stroke-Related Studies

Databases and Sites Searched for Stroke	
AJOT	PEDro
BJOT	PubMed
CINAHL	ScienceDirect
CJOT	www.rehabmeasures.com
OTSeeker	www.strokengine.ca

Table 3

Stroke-Related Studies Search Strategy and Results

Search key	Filters	Database	Hits	Date
Action Research Arm Test	Exact phrase	AJOT	13	10/26/2015
Action Research Arm Test OR ARAT	1995-present	BJOT	8	2/06/2016
(“Action Research Arm Test” OR ARAT) AND (psychometric OR reliab* OR valid* OR specificity OR sensitivity OR rasch)	Full-text, abstract available, 1995-2015, English	CINAHL	16	10/24/2015
Action Research Arm Test OR ARAT		CJOT	0	2/2/16
Action Research Arm Test OR ARAT		OT Seeker	0	10/26/2015
Action Research Arm Test		PEDro	97	10/24/2015
(("Action Research Arm Test" OR ARAT)) AND (psychometric OR reliabil* OR valid* OR specificity OR sensitivity OR rasch)	English, 1995-2015, full-text	PubMed	116	10/24/2015
“Action Research Arm Test” AND (psychometric OR reliab* OR valid* OR	Journal or review article, 1995-present, full-text	ScienceDirect	136	10/24/2015

clinimetric OR specificity OR sensitivity OR rasch)				
Arm Motor Ability Test OR AMAT OR Arm Mobility Arm Test	1995-present, assessment development and testing, adult,	AJOT	1	1/16/2016
Arm Motor Ability Test OR AMAT OR Arm Mobility Arm Test		BJOT	1	2/06/2016
(Arm Motor Ability Test OR Arm Mobility Arm Test OR AMAT) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)		CINAHL	9	10/25/2015
Arm Motor Ability Test OR AMAT		CJOT	0	2/2/16
Arm Motor Ability Test OR AMAT OR Arm Mobility Arm Test		OTSeeker	0	1/16/2016
Arm Mobility Arm Test		PEDro	31	10/25/2015
AMAT		PEDro	10	1/16/2016
Arm Motor Ability Test		PEDro	52	2/19/2016
(Arm Motor Ability Test OR Arm Mobility Arm Test OR AMAT) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)	Free full text, 1995-present, Humans, Clinical Trial, Comparative Study, Introductory Journal Article, Journal Article, Randomized Controlled Trial, Review, Systematic Reviews, Validation Studies, English,	PubMed	47	2/22/2016
(Arm Motor Ability Test) OR (Arm Mobility Arm Test) OR (AMAT) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)	1995-present, title/abstract/key terms	Science Direct	41	1/16/2016
“The Chedoke Arm and Hand Activity Inventory”		AJOT	5	10/24/2015
“The Chedoke Arm and Hand Activity Inventory”		BJOT	0	2/06/16

“The Chedoke Arm and Hand Activity Inventory”		CINAHL	19	10/24/2015
“The Chedoke Arm and Hand Activity Inventory”		CJOT	0	2/2/16
“The Chedoke Arm and Hand Activity Inventory”		OT Seeker	0	10/24/2015
“The Chedoke Arm and Hand Activity Inventory”		PEDro	0	10/24/2015
“The Chedoke Arm and Hand Activity Inventory”		PubMed	22	10/24/2015
“The Chedoke Arm and Hand Activity Inventory”		Science Direct	48	10/24/2015
Motor Evaluation Scale for Upper Extremity in Stroke Patients	Research article, hand and upper extremity, and 1995-present	AJOT	14	11/13/2015
Motor Evaluation Scale for Upper Extremity in Stroke Patients OR MESUPES		BJOT	0	2/06/2016
(Motor Evaluation Scale for Upper Extremity in Stroke Patients OR MESUPES) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)	1995-present	CINAHL	2	11/10/2015
Motor Evaluation Scale for Upper Extremity in Stroke Patients OR MESUPES		CJOT	0	2/06/2016
Motor Evaluation Scale for Upper Extremity in Stroke Patients		OT Seeker	0	11/13/2015
Motor Evaluation Scale for Upper Extremity in Stroke Patients		PEDro	8	2/06/2016
("Motor Evaluation Scale for Upper Extremity in Stroke Patients" OR "MESUPES") AND (psychometric* OR valid* OR reliab* OR	Comparative studies, journal articles, meta-analyses, reviews, systematic reviews, validation studies, full-text, 1995-present, humans	PubMed	65	11/13/2015

sensitivity OR specificity OR clinimetric* OR Rasch)				
(Motor Evaluation Scale for Upper Extremity in Stroke Patients OR MESUPES) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)	Abstract/title/key terms, journals, 1995-present	ScienceDirect	3	11/13/2015
"Streamlined Wolf Motor Test" OR S-WMFT		AJOT	0	11/12/2015
"Streamlined Wolf Motor Test" OR S-WMFT		BJOT	0	2/06/2016
("Streamlined Wolf Motor Function Test" OR S-WMFT) AND (reliab* OR valid* OR psychometric OR specificity OR sensitivity OR rasch)		CINAHL	6	11/12/2015
"Streamlined Wolf Motor Test" OR S-WMFT		CJOT	0	2/2/16
"Streamlined Wolf Motor Test" OR S-WMFT		OT Seeker	2	11/12/2015
"Streamlined Wolf Motor Test" OR S-WMFT		PEDro	0	11/12/2015
("Streamlined Wolf Motor Function Test" OR S-WMFT) AND (reliab* OR valid* OR psychometric OR specificity OR sensitivity OR rasch)		PubMed	5	11/12/2015
("Streamlined Wolf Motor Function Test" OR S-WMFT) AND (reliab* OR valid* OR psychometric OR specificity OR sensitivity OR rasch)	In journals, 1995-present, articles or review articles	ScienceDirect	92	11/12/2015

Inclusion/ Exclusion of Stroke articles.

Inclusion criteria for stroke articles:

- Addressed upper extremity motor control
- Contained psychometric data
- Published within the past 20 years

Exclusion criteria for stroke articles:

- Did not address target population
- Focused on an intervention
- Published in a language other than English
- Not accessible to the researchers through either the University of Puget Sound or University of Washington school databases or inter-library loan systems
- Addressed an alternate version of the assessment published at an earlier date

After analysis of the articles following the exclusion criteria, it was decided that articles that compare the assessment to one or more assessments not within this scope of study would be excluded as well. By adding this additional exclusion criteria, it was ensured that all the articles contained only relevant information regarding the assessments chosen.

Action Research Arm Test (ARAT). Searching the databases listed above resulted in 386 articles. Of those articles, 371 were excluded: 276 focused on interventions, seven were not in English, 59 analyzed different assessments, 10 were duplicates from other databases, six targeted a different population, three were not available in full text, one contained no psychometric data and nine compared the ARAT to assessments not of interest. A total of 15 articles regarding the ARAT were included in the CAT.

Arm Motor Ability Test (AMAT). Searching the databases listed above resulted in 192 articles. Of those articles, 187 were excluded: 167 focused on interventions, 15 analyzed different assessments, two were duplicates from other databases, one targeted a different population, one was unavailable in full

text and one compared the AMAT to an assessment not of interest. A total of five articles regarding the AMAT were included in the CAT.

Chedoke Arm and Hand Activity Inventory (CAHAI). Searching the databases listed above resulted in 94 articles. Of those articles, 87 were excluded: four were duplicates from other databases, three were unavailable in full text, and 80 were not relevant to the assessment. A total of seven articles regarding the CAHAI were included in the CAT.

Motor Evaluation Scale for Upper Extremity in Stroke Patients (MESUPES). Searching the databases listed above resulted in 92 articles. Of those articles, 90 were excluded: 35 focused on interventions, 52 analyzed different assessments, two were duplicates from other databases, and 1 targeted a different population. A total of two articles regarding the MESUPES were included in the CAT.

Streamlined Wolf Motor Function Test (S-WMFT). Searching the databases listed above resulted in 105 articles. Of those articles, 101 were excluded: 77 focused on interventions, six were not in English, four analyzed other assessments, six were duplicates from other databases, three contained no psychometric data, one compared the S-WMFT to an assessment not of interest, and four were a different version of the assessment. A total of four articles regarding the S-WMFT were included in the CAT.

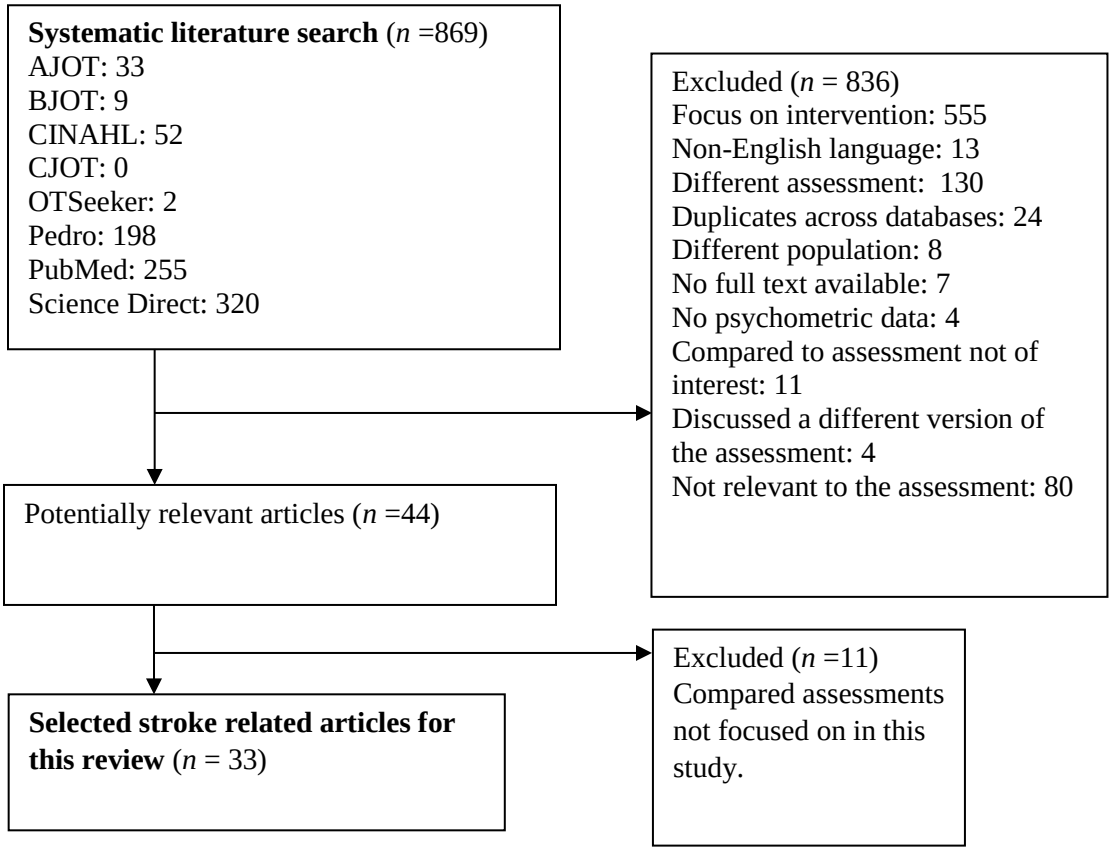

Figure 2: Flow chart of the identification of studies on stroke assessments for inclusion in this critical appraisal

Search Strategy for PD articles. A search using the exact search string "parkinson" AND "upper extremity" AND "assessment" through PubMed returned a recent systematic review of upper extremity measures specifically for patients with Parkinson's disease done by Proud et al. (2015). Articles from inception to November 2013 were included in that review. Repeating the search strategy and applying a date filter of November 2013 to February 2016, no other articles were found. Proud et al.'s (2015) systematic review was utilized because the purpose of their review was to discover measurement tools used for upper limb evaluation in people with PD. Because the Rush Dyskinesia Rating Scale was added to the list of recommended PD assessments after searching Rehab Measures, a search of databases for the assessment was conducted (See Table 6).

Parkinson's Disease Related Articles Search Process

Table 4

Search Terms for Parkinson's Disease-Related Studies

Key Terms	**Synonyms**	**Alternate Spellings**
Parkinson's Disease		Parkinson*, PD
Psychometric	clinimetric, reliable, valid, sensitivity, specificity	reliab*, valid*
Rush Dyskinesia Rating Scale		
Upper Extremity	Arm, hand, upper limb,	UE

Table 5

Databases Searched for Parkinson's Disease-Related Studies

Databases and Sites Searched for Parkinson's Disease	
AJOT	PEDro
BJOT	PsychINFO
CINAHL	PubMed
CJOT	ScienceDirect
Medline	www.rehabmeasures.com
OTSeeker	

Table 6

Parkinson's Disease-Related Studies

(Rush Dyskinesia Rating Scale) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)		AJOT	0	1/16/2016
(Rush Dyskinesia Rating Scale) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)		BJOT	0	1/16/2016
(Rush Dyskinesia Rating Scale) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)	1995-present, linked full text, academic journals	CINAHL	0	1/16/2016
(Rush Dyskinesia Rating Scale) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)		CJOT	0	3/2/16
(Rush Dyskinesia Rating Scale) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)		OTSeeker	0	1/16/2016
(Rush Dyskinesia Rating Scale) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)		PEDro	0	1/16/2016
Rush Dyskinesia Rating Scale	Full text, 1995-present, Humans, Clinical Trial, Comparative Study, Introductory Journal	PubMed	31	1/16/2016

	Article, Journal Article, Randomized Controlled Trial, Review, Systematic Reviews, Validation Studies,			
(Rush Dyskinesia Rating Scale) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)	Full text, 1995-present, Humans, Clinical Trial, Comparative Study, Introductory Journal Article, Journal Article, Randomized Controlled Trial, Review, Systematic Reviews, Validation Studies,	PubMed	10	1/16/2016
(Rush Dyskinesia Rating Scale) AND (psychometric* OR valid* OR reliab* OR sensitivity OR specificity OR clinimetric* OR Rasch)	1995-present, Abstract/Title/Keywords, Journals,	ScienceDirect	1	1/16/2016
(Parkinson OR "Parkinson* disease") AND ("upper limb" OR "upper extremity" OR arm OR hand) AND (measure* OR assess* OR outcome OR tool OR instrument* OR dexterity OR disability OR ADL)	English, research article, 2013-present	CINAHL	24	2/06/2016
(Parkinson OR "Parkinson* disease") AND ("upper limb" OR "upper extremity" OR arm OR hand) AND (measure* OR assess* OR outcome OR tool OR instrument* OR dexterity OR disability OR ADL)	2013-present, linked full text, academic journals	MEDLINE	32	2/06/2016
Parkinson's Disease Upper Extremity Assessment		OTSeeker	0	2/06/2016
(Parkinson's Disease OR PD) AND (upper extremity OR UE) AND (clinimetric* OR reliab*		OTSeeker	0	2/06/2016

OR valid* OR sensitivity OR specificity)				
Parkinson's Disease AND Upper Extremity AND Assessment		PEDro	0	2/06/2016
(Parkinson OR "Parkinson* disease") AND ("upper limb" OR "upper extremity" OR arm OR hand) AND (measure* OR assess* OR outcome OR tool OR instrument* OR dexterity OR disability OR ADL)	2013-present, linked full text, academic journals	PsychINFO	14	2/06/2016
"parkinson" AND "upper extremity" AND "assessment"		PubMed	1 (PD systematic review)	10/24/2015
(Parkinson OR "Parkinson* disease") AND ("upper limb" OR "upper extremity" OR arm OR hand) AND (measure* OR assess* OR outcome OR tool OR instrument* OR dexterity OR disability OR ADL)	Full text, 2013-present, Humans, Clinical Trial, Comparative Study, Introductory Journal Article, Journal Article, Randomized Controlled Trial, Review, Systematic Reviews, Validation Studies,	PubMed	36	2/06/2016

Inclusion/ Exclusion Criteria for PD articles.

Inclusion criteria for articles in Proud et al. (2015) systematic review:

- Participants were diagnosed with idiopathic PD
- Evaluated upper limb impairments activity limitations, and/or participation restrictions
- Assessments evaluated the effects of treatment or disease progression
- Focused on a measurement method that is replicable within clinical practice

Exclusion criteria for articles in Proud et al. systematic review:

- Not published in English
- Did not provide detailed information regarding the assessment or assessment protocol

- Utilized lab based assessments, such as kinematic analysis
- Contained obsolete commercially produced tools
- Letters, editorial, literature reviews and conference abstracts

Inclusion criteria for Rush Dyskinesia Rating Scale articles:

- Addressed upper extremity functional outcomes
- Contained psychometric data
- Published within the past 20 years

Exclusion criteria for Rush Dyskinesia Rating Scale articles:

- Did not address our target population
- Focused on an intervention
- Published in a language other than English
- Not accessible to the researchers through either the University of Puget Sound or University of Washington school databases or inter-library loan systems
- Addressed an alternate version of the assessment published at an earlier date

Rush Dyskinesia Rating Scale (RDRS). Searching the databases listed above resulted in 42 articles. Of those, 40 articles were excluded: 11 focused on other assessments, 26 focused on interventions, one was not available in full-text, and two were duplicates from other databases. Two articles from this search, in addition to the systematic review by Proud et al. (2015), were included in the CAT.

Other Assessments Appropriate for Patients with PD Search. Repeating the search strategy of Proud et al. (2015) to find articles on assessments appropriate for use with patients with Parkinson's disease was largely unsuccessful. Using six of the nine databases we had access to and following the authors' search strategy, there were 106 hits. All 106 were excluded: 100 articles were excluded because they focused on interventions and neurology topics, three analyzed different assessments, one analyzed a non-standardized assessment, one did not include psychometric data, and one was a duplicate (the systematic review included in this research).

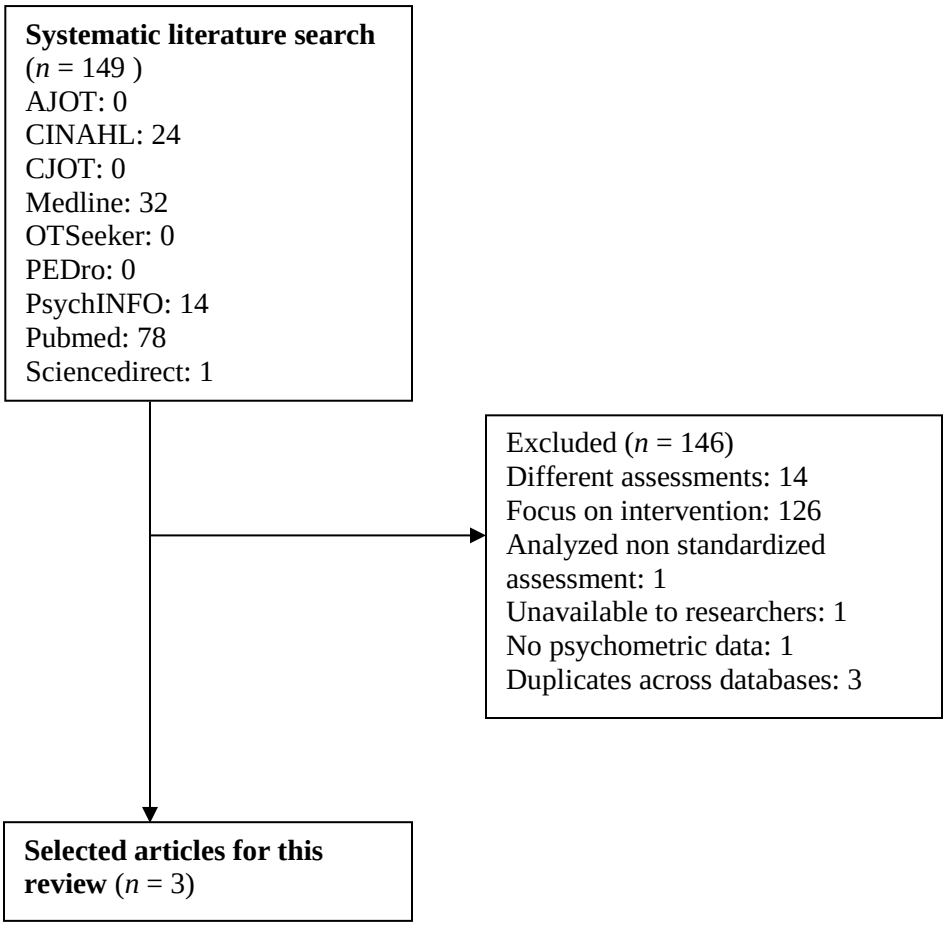

Figure 3: Flow chart of the identification of studies on Parkinson's disease assessments for inclusion in this critical appraisal

RESULTS

Stroke Assessments

The Action Research Arm Test Results

Author, Year	Study Objectives	Study Design/ Level of Evidence	Participants: Sample Size, Description Inclusion and Exclusion Criteria	Interventions & Outcome Measures	Summary of Results	Study Limitations
Chen, Lin, Wu, & Chen (2012)	Validate the internal construct and predictive validity of the ARAT.	Secondary study, association/ correlation; Pyramid level: D2, AOTA level: IV	Pts with stroke drawn from ongoing research into two interventions; *N* = 191 Inclusion: first time stroke, Brunnstrom ≥ 2, MMSE > 21, MAS < 2, no severe medical issues, no joint pain, no balance-related safety issues. Mean age 55.17 yrs, mean time since stroke 17.19 mo.	Rasch analysis for construct validity; post-tx score correlation between ARAT, WFMT, MAL, SIS for predictive validity; pts randomly assigned to receive BAT, CIMT, or control intervention.	Predictive validity fair with the SIS-Physical (p = 0.45); moderate with MAL-AOU (p = 0.62), SIS - hand function (p = 0.58), and WMFT-TIME (p = -0.66); good with the WMFT-FAS (p = 0.76). Construct validity was disordered; authors recommend moving from a 4-point to a 3-point scale (collapse scores 0 and 1).	Limited to stroke survivors (both chronic and sub-acute), no cognitive limitations, motor control return present in affected limb. All participants only had mild to moderately impaired UE function with a mean time post stroke of 17.19 mo.
Hsieh, Hsueh, Chiang, & Lin (1998)	Verify the inter-rater reliability and validity of the ARAT.	Prospective association/ correlation; Pyramid level: D2, AOTA level: IV	*N* = 50. Inclusion: stroke pts, consecutive hospital admissions with stroke and the ability to follow verbal commands; mean age	ARAT administered by 3 different therapists on 3 days to find inter-rater reliability;	ICC of 0.98 (95% confidence interval 0.97-0.99, F = 178.3, p < 0.0001); ANOVA found systematic bias on 3 subscales (grip, pinch,	Older population, no pts with global aphasia, relatively low

			65 yrs, mean time after stroke 55 days.	concurrent validity found by comparison with the MI, MAS, and UEMMAC.	grasp); correlation r = 0.96 with the MAS, r = 0.87 with the MI, and r = 0.94 with UEMMAC. Results supported validity of the ARAT as a measure of UE function.	sample size, inter-rater reliability established with only 3 therapists.
Hsieh, Wu, Lin, Chang, Chen, & Liu (2009)	Investigate and compare responsiveness and validity of the ARAT, WMFT, and the FMA.	Association/ correlation; Pyramid level: D2, AOTA level: IV; data pulled from RCT (E2, I)	N = 57. Pts with chronic stroke recruited from 3 medical centers. Inclusion: first time stroke > 6 mo past, Brunnstrom stage ≥ 3, MAS ≤ 2.5, MMSE ≥ 24, no confounding medical conditions. Measurements taken pre- and post-tx; mean age 54.56 yrs, mean time since stroke 12.98 mo.	Subjects receiving CIMT, BAT or conventional therapy for three wks; outcome measures compared to FIM scores as external criterion; responsiveness tested with Wilcoxon matched pairs, SRM; validity found with Spearman correlation coefficient.	Responsiveness of FMA higher than ARAT (difference in SRM = 0.47) and WMFT-TIME (difference in SRM = 1.04), but not WMFT-FAS (difference in SRM = 0.12); construct validity: at pre-tx, ARAT had good correlation with other outcome measures (p = 0.63-0.77); at post-tx, ARAT had moderate to good correlation (p = 0.58-0.74); ARAT had low predictive validity with FIM scores. FMA has sound responsiveness and validity and is good for predicting a pts UE functioning post-tx. Authors recommend use of the FMA over ARAT or WMFT for assessing pts undergoing stoke rehab.	Relatively low sample size, limited to mild to moderate chronic stroke with high cognition.
Koh, Hsueh,	Determine if	Association/	N = 351. Recruited	Study did not	Unidimensionality	Study limited

Wang, Sheu, Yu, Wang, & Hsieh (2006)	ARAT is unidimensional, and if test items can be made interval with Rasch analysis.	correlation; Pyramid level: D2, AOTA level: IV	from 5 in- and outpatient rehab departments, invited to participate if they had dx of stroke, ability to follow instructions, and absence of other major diseases or impairments. Median age 63 yrs, mean time since stroke 12.5 mo.	include intervention; ARAT administered by same PT to all participants; Mokken analysis to determine dimensionality, Rasch analysis to determine if scores can be made interval.	established: scalability coefficient of 19-item ARAT $H = 0.95$; one item (pinch ball bearing 3rd finger and thumb) over Crit value benchmark of 80, when removed, scalability coefficient of 18-item ARAT $H = 0.95$, Pmatrix < 57, Restscore <15, indicating that items of 18-item ARAT difficulty order is the same for all pts; only 4 items of the test (grasp ball, grasp block 5cm^3, grasp block 2.5 cm^3, grip tube 1 cm^3) fit Rasch analysis, thus the authors recommend that clinicians use the 19 items of the ARAT as a whole, and not divide scores into the four subscales.	to chronic stroke pts with no cognitive impairments and no comorbidities. Sensory impairments were not addressed. Assessment administered by a PT.
Lang, Edwards, Birkenmeier, & Dromerick (2008)	Estimate minimal clinically important difference in ARAT, WMFT, and MAL.	Association/ correlation; Pyramid level: D2, AOTA level: IV; data pulled from pilot RCT (E2, I)	$N = 52$.Pts from stroke research registries of VECTOR. Inclusion: stroke within 28 days, NIHSS motor arm items score 1-3, some voluntary UE motor return (ability to move jts against gravity), NIHSS consciousness	Randomized into 1 of 3 tx groups: 2 hrs conventional tx, 2 hrs shaping + 2 hrs CIMT, 3 hrs shaping + constraint 90% waking hrs; pre- and post-testing done by blinded	MCID if dominant hand affected for ARAT: 12 raw value, 21 percent of total, 0.78 effect size. ARAT non-dominant hand: 17 raw score, 30 percent of total, 1.10 effect size. MCID scores of 16-30% were considered important to	Limited to acute stroke; unable to find MCID in all instances; older population, intact cognition, in-pt setting

			score 0-1, SBMOC ≥ 19, ability to follow 2-step commands, no prior UE injury. Exclusion: no fluctuation in mental state, no UN, life expectancy > 1 yr. Mean age 64 ± 14 yrs, mean time since stroke 9.5 ± 4.5 days.	assessor; ARAT, WMFT, and MAL measured at start and after 14 days; MCID calculated for each measure using SPSS. Grip strength and composite UE strength also tested, but are not applicable to this study so results are not included.	pts. Thus, the ARAT established a satisfactory level of MCID for the pts in this study.	may not generalize to SNF.
Lang, Wagner, Dromerick, & Edwards (2006)	Examine responsiveness and validity of the ARAT in pts with acute stroke.	Association/ correlation; Pyramid level: D2, AOTA level: IV, data pulled from pilot RCT (E2, I)	*N* = 50, all subjects enrolled in VECTOR; NIHSS arm motor score 1-2, ability to move jts against gravity, ability to follow 2-step directions, no prior UE injury, (see above exclusion criteria). Mean age 63.7 ± 13.6 yrs, mean time since stroke 9.5 ± 4.5 days.	Randomized into groups described above; measures given by trained, blinded personnel at baseline, after 14 days of tx, and at 90 days for follow-up; stroke severity measured with NIHSS, light touch with Semmes-Weinstein monofilaments, pain with VAS, spasticity with MAS, FIM for disability measure (via phone).	Responsiveness: single population effect size from day 0-14 = 1.018 (large), from day 0-90 = 1.390 (large); ARAT responsive to change during first wks and mo. following stroke. MCID is 10% (6 points on a 57-point scale) indicating ARAT can detect smallest meaningful change. Construct validity (results given for day 0, day 14, and day 90): correlation with age *p* = -0.16, -0.44, -0.29 (moderately related at day 14); with stroke severity *p* = -0.15, -0.24, -0.29 (minimally related);	Sample included pts with mild-to-moderate hemiparesis in acute post-stroke, older population, intact cognition, no sensory impairments, some voluntary motor control return; in-pt setting may not generalize to SNF; tx group intervention

					Correlation between strength and ARAT moderate ($p \approx 0.5$), poor with light touch, pain, spasticity ($p \approx 0.0$, -0.1, -0.5); ARAT and FIM scores not well correlated (measurements varied over 90 days, no numerical results given). Near-perfect scores on the ARAT does not necessarily signify no disability, clinicians may consider using it with other disability measures.	time not equal; FIM scored over telephone; numerical results not given for all tests.
Li, Lin, Wang, Wu, Huang, & Ouyang (2012)	Examine the demographics and 3 measures of motor function in determining outcomes of ADL after distribution of dCIT.	Descriptive correlational; Pyramid level: D2, AOTA level: IV	*N*= 69. Inclusion: dx of stroke >1 mo. prior to study, follow verbal instructions, Brunnstrom stage 3, Modified Ashworth Scale 2 in any jt. Mean time since stroke 16.68 mo. Mean age 56.56. 49 female, 29 male.	Pts performed dCIT training for 3 wks. Pts evaluated at baseline and 3 wks post dCIT intervention completion. Evaluation measures used: ARAT, FMA, WMFT, MAL, NEADL, and SIS. Administered by an occupational therapist.	ARAT and FMA are better at predicting ADL/IADL outcomes than WMFT. ARAT grasp-grip-pinch movement score is most suitable predictor for self-report improvement in ADL/IADL. FMA is best predictor of participation and UE functional use. Age was the only significant demographic predictor with less reported improvement in older adults. Predictive validity: with NEADL 0.273, with SIS-	Study did not consider predictive ability of assessments for clients with diverse cognitive impairment, family support, medication, mental health, or emotional status. Results may not apply to pts who

					ADL/IADL 0.360, with MAL 0.443-0.454	receive other interventions.
Lin, Hsu, Sheu, Wu, Lin, Chen, & Hsieh (2009)	To compare psychometric properties of UE-FMA, UE-STREAM, ARAT and WMFT.	Descriptive correlational; Pyramid level: D2, AOTA level: IV	*N* = 35. Participants included in validity analysis. Inclusion: dx of first stroke, onset within 2 wks before hospital admission, able to follow instructions. *N* = 30 participants included in interrater reliability analysis. *N*= 30 participants included in test-retest reliability analysis. Inclusion: >1 yr post-stroke, stable medical condition, able to follow instructions; follow-up at 6 months.	Two stages: 1) validity, interrater reliability, and responsiveness determined in pts consecutively admitted to a neurology department with sub-acute stroke; 2) test-retest reliability investigated with pts with chronic stroke recruited independently of first stage. Physical therapist administered 4 UE assessments and the BI at 14, 30, 90, and 180 days post-stroke to determine validity and responsiveness. Two physical therapists administered the 4 assessments 14 days post-STROKE to determine interrater reliability. One therapist	UE-STREAM, ARAT, and WMFT had significant floor effects at 14 days post-stroke and ceiling effects at 30, 90, and 180 days. UE-FM floor and ceiling effects not notable and it is more discriminative for individuals with either very poor or very good motor function. ARAT was most responsive, with better responsiveness later in recovery (effect size 0.49-0.79). Evidence for good clinical utility, test-retest reliability (ICC 0.99), concurrent and predictive validity across all 4 assessments (0.81-0.97). Findings suggest changes of >3, 4, 6, and 12 points for UE-STREAM, ARAT, UE-FM, and WMFT, respectively, are unlikely to be due to measurement error or chance; authors suggest that the minimal detectable change for the	Small sample size; considering severity or type of stroke in analysis was not possible. Age of participants lower than typical age of stroke onset in Taiwanese people, meaning Limited generalizatio n of results to clinician setting; set in Taiwan, no cognitive impairments included, administered by PTs, did not discuss level of training for therapists to achieve IRR levels.

				administered 4 assessments twice, 1 wk apart to measure test-retest reliability. Testing was random and counterbalanced.	ARAT is satisfactory.	Did not use timed aspect of WMFT.
Ng, Leung, & Fong (2008)	To study the ARAT, WMFT and MAL in pts with stroke and UE impairment.	Descriptive; Pyramid level: D2, AOTA level: IV	*N*= 12. Seven participants in lower functioning group and 5 in higher functioning group. Participants attended a community center in Hong Kong. Mean time since stroke 69.3 mo. Mean age 57.3. Inclusion: time since stroke >6 mo., able to sit for 30 min., no receptive language problems, can follow 1-2 step commands.	ARAT, WMFT, MAL and FTHUE-HK were administered among participants one time. ARAT and MAL were administered using a standardized approach. Only 15 functional tasks of the WMFT were tested. Used SPSS and Spearman's correlation coefficients to describe relationships between measures. Signigificance level: p<.05	Participants were stratified into two groups based on results of FTHUE-HK. Lower scores on the FTHUE-HK were good predictors for lower scores on the ARAT and WMFT. FTHUE-HK was highly statistically correlated to the WMFT and ARAT (rho = 0.92, $p < 0.001$). ARAT tests functional tasks and has hierarchical items such that subtests can be skipped, potentially saving time. ARAT prone to high floor affect; had a difficult time detecting function in pts with severe impairment but not the case with mild impairment. WMFT shown to have both floor and ceiling affects. ARAT more useful with high functioning pts,	Pilot study: mean age is low relative to typical age of stroke. Large variation in time since stroke. Participants are >6 mo. post-stroke which is not representative of a typical subacute setting. Small convenience sample meaning limited generalizability. FTHUE-HK designed to be culturally relevant for Hong Kong

					WMFT useful with low functioning pts.	residents.
Nijland, van Wegen, Verbunt, van Wijk, van Kordelaar, & Kwakkel (2010)	To study concurrent validity between ARAT and WMFT. To compare their reproducibility, internal consistency and floor and ceiling effects.	Association/ correlation; Pyramid level: D2, AOTA level: IV	*N*= 40. Pts from 2 rehab centers in the Netherlands. Dx of stroke. Mean age 60 yo. Median time since stroke .41 yrs. Inclusion: hemiparesis of UE with some voluntary control, MMSE score > 22, no orthopedic UE limitations.	Pts at first center participated in reproducibility testing, both therapists administered tests over one wk in random order; data from both centers used for internal consistency. Standardized method of scoring developed by Yozbatarin et al, was used.	Good inter- and intra-observer reliability for ARAT (0.92-0.97) and WMFT. More conflict of scores within observer for the WMFT than ARAT. Cronbach's alpha score of 0.98 or higher for ARAT and WMFT show high internal consistency. No floor and ceiling effects for ARAT or WMFT. ARAT concurrent validity with WMFT (rho = 0.86.)	Modest sample size. Only included pts with mild-moderate stroke.
Nordin, Murphy, & Danielsson (2014)	To determine the intra- and inter-rater reliability of the (ARAT) at the item level after stroke.	Descriptive; Pyramid level: D2, AOTA level: IV	*N*= 35. Pts with impaired UE function post stroke. Median age 62 yrs, median time post-stroke 22 mo. 8 female, 27 male. Exclusion: absence of active movement in the affected arm; other dx affecting UE, and incomprehension of Swedish language.	Pts were assessed using the ARAT 2xs in one day. Test was simultaneously administered and scored by two physiotherapists both times for each pt. A rank-based statistical method for paired ordinal data was used.	Satisfactory intra- and inter-rater agreement achieved for all items on the ARAT (percentage agreement 89-94%) except item 19, which was below satisfactory. Non negligible systematic disagreements on 6 items were found within and between raters. No disagreement due to random variance within or between raters. ARAT found to be highly reliable.	Only performed by two different therapists. Administered in Swedish language. All participants only had mild to moderately impaired UE function with a mean time post stroke of 22 mo. Clinicians

						had up to one year experience using ARAT and up to 30 yrs therapy experience.
Van Delden, Peper, Beek, & Kwakkel (2013)	Explore match and mis-match between and objective measure (ARAT) and subjective measures (MAL and SIS-Hand) of improvement post-stroke; which participant determinants factored into matches	Association/ correlation; Pyramid level: D2, AOTA level: IV; data pulled from early stages of ULTRA-stroke RCT (E2, I)	*N* = 39. Inclusion: First stroke, UE paresis (with 10° wrist ext, thumb abd/ext, and ext in at least 2 digits), ARAT < 53, age 18-80, consenting, motivated, no orthopedic limitations, UCO ≥ 3, MMSE ≥ 23; mean age 61.4, mean time since stroke 9.3 wks.	Participants randomly assigned to receive CIMT, BATRAC, or conventional therapy for 3 wks, follow-up after 6 wks; mood measured using the SIS-Emotion domain.	Probability = 0.67 that clinically meaningful improvement on ARAT matched with meaningful change on MAL-QOM for all determinants; probability = 0.83 for match between ARAT improvement and SIS-Hand improvement when education low (less than bachelor's degree) and mood above normative score.	Small sample, RCT data pulled from incomplete, no baseline data by intervention, ULTRA-stroke inclusion criteria limits generalizability. All participants only had mild to moderately impaired UE.
Van der Lee, De Groot, Beckerman, Wagenaar, Lankhorst & Bouter (2001)	Discover intra- and inter-rater reliability of the ARAT. To assess its ability to detect a MCID. To discover less reliable test	Descriptive correlational; Pyramid level: D2, AOTA level: IV	*N*= 20, out of an RCT sample of *N*= 66. Pts were arranged by their intake ARA score and every 3rd pt was selected to be included in the study. Pts with chronic stroke, median age 62yrs, median	ARAT administered 2xs at baseline, 2xs during 2 wk therapy intervention phase, and 4 times post-intervention. Every other administration was	ICC and Spearman's rho > 0.98, demonstrating high intra- and interrater reliability of the ARAT. Some items and subtests are easier to rate than others. For instance, it was difficult to distinguish between a	Time limit on the difficult-to-rate items is shorter than other items and they failed to ensure each pt's back

	items.		time since stroke onset 3.6yr. Inclusion: hx of stroke, no severe cognitive impairments or aphasia, min. of 20 degrees of active extension in the wrist and 10 degrees of finger extension.	videotaped. Administrator rescored pts via rewatching the video 4-27 mo. post administration, and again 4-6 wks later. Pts were also scored via videotape by a second rater. Two different ways of scoring were analyzed: 1. By adding the score of all 19 tasks. 2. By breaking all tasks into four subgroup and testing the hardest task in the subgroup first. If pt passes first task, move on to next subgroup, if not, have pt do easiest task in that subgroup. All 19 tasks were completed by all pts, but were analyzed as if it was administered using the 2nd protocol as well.	score of 2 and 3 on the "hand to mouth" item. ARAT is capable of detecting a difference of at least 5.7 points, or 10% its maximum score of 57 points. That MCID is considered clinically relevant.	remained against the backrest of the chair when rating. The sample did not include low-functioning pts. Limited to chronic stroke. All participants only had mild to moderately impaired UE.
Van der	Determine the	Association/	*N* = 63. Inclusion: hx	Participants	Unidimensionality	Excluded

Lee, Roorda, Beckerman, Lankhorst, & Bouter (2002)	dimensionality of the ARAT, which method of the test (Lyle's original, all 19 items, 15 items) is most efficient and effective.	correlation; Pyramid level: D2, AOTA level: IV; data pulled from RCT (E2, I)	of single stroke at least one yr prior, 20° ext in wrist, 10° ext in fingers, ARAT < 51, able to walk without aids, no severe aphasia, MMSE ≥ 22. Median age 61 yrs, median yrs since stroke 3.0.	involved in forced use tx. *n*(intervention group)= 31, *n*(control group)= 32.	established: scalability coefficient of 19-item ARAT *H* = 0.79, reliability coefficient rho = 0.98 (high internal consistency); when 4 of 6 pinch subtest items (items 1, 3, 4, 5) were removed, the scalability coefficient of the 15-item ARAT *H* = 0.83 with rho = 0.97, representing a unidimensional hierarchical scale; responsiveness ratio higher for 19- and 15-item versions over Lyle's original structure. These results show that the ARAT can be altered by omitting 4 test items and still be statistically relevant This can decrease the time it takes to administer.	aphasia, sensory and cognitive impairments, motor return present in wrist and hand, limited to chronic stroke, used non-standardized time limit on items.
Yozbatiron, Der-Yeghiain, & Cramer (2008)	Present standardized approach to administer the ARAT, provide reliability and validity data for this approach.	Association/ correlation; Pyramid level: D2, AOTA level: IV	*N* = 12. Inclusion criteria: chronic stroke (> 3 mo prior), moderate R hemiparesis (> 10° ROM at MCP jt, R 9-HPT < 75% of L), age > 18 yrs, R hand dominant. Mean age 61 ± 15 yrs, mean time	Study did not include intervention; interrater reliability found between 2 therapists over 9 participants, intrarater by 8 participants over 1 week by 1	Interrater ICC = 0.9986, intrarater ICC = 0.99 (very high for both); high correlation between ARAT and UE-FMA (r = 0.94, P < 0.01); authors outline standardized method for administration in the appendix, including	Limited to chronic stroke, no information on where participants recruited from. All participants only had

			since stroke 34 ± 59 mo.	therapist, validity by comparison with the UE-FMA. Provides the most detailed manual with instructions on how to administer and score the ARAT in order to increase reliability.	specific equipment and scoring guidelines.	moderately impaired UE.

Abbreviations Key: 9-HPT – 9-Hole Peg Test, abd – abduction, ARAT - Action Research Arm Test, BAT - bilateral arm training, BATRAC – bilateral arm training with auditory cueing, BI – Barthel Index, CIMT - Constraint Induced Movement Therapy, dCIT – distributed constraint-induced therapy, dx - diagnosis, ext - extension, FMA – Fugl-Meyer Assessment, FTHUE-HK – the Functional Test for the Hemiplegic Upper Extremity Hong Kong version, hrs – hours, hx – history, ICC – intraclass correlation coefficient, jts – joints, L - left, MAL - Motor Activity Log, MAL-AOU – Motor Activity Log amount of use, MAL-QOM – Motor Activity Log quality of movement, MAS - Modified Ashworth Scale, MCID - minimal clinically important difference, MCP - metacarpophalangeal, MI – Motricity Index, MMSE - Mini-Mental State Examination, mo – month, NEADL – Nottingham Extended Activities of Daily Living questionnaire, NIHSS – National Institutes of Health Stroke Scale, pt – patient, PT – physical therapist, R - right, RCT – randomized controlled trial, ROM – range of motion, SBMOC – Short Blessed Memory Orientation and Concentration test, SIS - Stroke Impact Scale, SIS-Hand – Hand domain of the Stroke Impact Scale, SNF – skilled nursing facility, SRM – Standardized response mean, tx - treatment, UEMMAC - upper extremity Modified Motor Assessment Chart, UE - upper extremity, UE-FMA – upper extremity section of FMA, UE-STREAM – upper extremity portion of Stroke Rehabilitation Assessment of Movement, ULTRA-stroke – Upper Limb Training after stroke trial, UN – unilateral neglect, VAS – visual analog scale, VECTOR - Very Early Constraint-induced Therapy for Recovery of Stroke, wks - weeks, WMFT - Wolf Motor Function Test, WMFT-FAS - functional ability scale of the WMFT, WMFT-TIME - performance time on WMFT, UCO – Ultrecht Communication Observation.

The Arm Motor Ability Test Results

Author, Year	Study Objectives	Study Design/ Level of Evidence	Participants: Sample Size, Description Inclusion and Exclusion Criteria	Interventions & Outcome Measures	Summary of Results	Study Limitations
Kopp, Kunkel, Herta Flor, Platz, Mauritz, Gresser, McCulloch, & Taub (1997)	Replicate AMAT reliability and sensitivity findings.	Descriptive, correlational; Pyramid level: D2, AOTA level: IV	*N* = 33. 12 female, 21 male. Primarily in the subacute phase of stroke (median post-stroke days: 43). Median age 66 yrs. Inclusion: >20 degrees wrist flexion, 10 degrees at MC and IP joints, MI (UE only) score 66-99.9, MMSE score>19.	Randomly assigned to one of two groups. All received two administrations of AMAT. Four components of original AMAT not included. Administrations were either 1 or 2 weeks apart, depending on the group. 13/17 timed, compound ADL tasks were evaluated. Measures: AMAT (Functional ability, quality of movement, and performance time) & MI.	Interrater reliabilities for the AMAT were high (Kappas: 0.68-0.77), Spearman coefficients of 0.95 or above). Test-retest reliability: 0.93-0.99. Homogeneity of AMAT very good. Internal consistency close to unity. Concurrent validity with UE portion of MI is satisfactory (r=0.61), but AMAT is internally consistent whereas MI is not. Sensitivity to change in pts' motor control was determined by comparing patients' test-retest data of 1-week and 2-week groups. Determined test could be shortened without affecting reliability or sensitivity.	Complete AMAT was not used in research. AMAT was compared to an assessment with mediocre internal consistency. Therefore, the concurrent validity cannot be derived from this study alone. Did not include low physical and/or low cognitive functioning pts.

O'Dell, Kim, Rivera, Fieo, Christos, Polistena, Fitzgerald, & Gorga (2013)	Explore psychometrics of a 9-item version of AMAT, to determine appropriateness for use with pts with stroke and severe impairments as sequelae.	Descriptive, correlational; Pyramid level: D2, AOTA level: IV	N = 32. 9 female, 23 male. 35-85 yrs old. 88% Right-handed with post-stroke range 0.8-25.2 yrs. Inclusion: ≥6 mo. post-stroke, able to follow directions, adequate vision, > 3 mo. since rehab tx or botulinum injection, adequate PROM to participate in robotics therapy, MM strength 1-⅘ for all UE joints.	WMFT, ARAT, FMA, and AMAT-9 assessments administered by single researcher at baseline and following 12 weeks of robotics treatments. SIS was completed by participants at home with family or in the clinic with a researcher. The assessment scores were compared through statistical analysis methods.	Participants demonstrated severe functional deficits overall. Reliability of AMAT-9 was good. Correlation coefficients between AMAT-9 and ARAT, FMA, and WMFT were significant and identical (0.78-0.79, $p<0.001$). AMAT demonstrated intermediate responsiveness compared to other assessments. AMAT considers compensatory behaviors unlike BI and FIM. Unidimensionality of AMAT-9 supported by independent t-tests (6.25% fell outside 1.96 +/- range). Characteristics are also consistent with a Rasch model. Rasch analysis demonstrated hierarchical item difficulty.	Participants not randomly selected. Participants were primarily young motivated individuals interested in participating after other treatment options were exhausted. Most participants had moderate to severe UE functional limitations. Assessor was not blinded to the study design or hypotheses. Scales were not randomly administered. Participants may not reflect the population typically seen in a subacute setting (>6 mo. post-stroke).
O'Dell, Kim, Finnen, &	After reviewing 4 studies the researchers	Literature Review; Pyramid	An unspecified number of articles were included.	Descriptive information about the content	Test is easy to administer, inexpensive, and uses	Information about the quality of articles was not included; therefore,

Polistena (2011) Including studies by: McCulloch, 1988 Kopp, 1997 Chae, 2003 Daly, 2005	provided an overview of current use of the AMAT including content evolution and provide recommendations for future use.	level: D1, AOTA level: I	Inclusion: Studies that addressed clients with stroke and considered psychometrics of AMAT and/or that included AMAT as an outcome measure.	evolution and psychometrics of the AMAT alone, including comparisons between different versions of the AMAT. Comparison of the psychometrics and outcomes of the AMAT, BI, FMA, SIS, FIM, and WMFT.	everyday task objects. The administrator must be skilled in multitasking in order to score. The test allows for bilateral arm use, so even pts with little hand function should have some success. Strong psychometrics for full and shortened version: test-retest and interrater reliability coefficients: 0.68-1.0; ICC: 0.94-0.97. Significant correlations between AMAT and other assessment scores (e.g. 0.75-0.90 with FMA). Sensitivity and responsiveness are not well established. FA should be only domain assessed since Kopp et al, found psychometric redundancy between FA and QOM, and Chae et al. found significant floor and ceiling affects for performance time and AMAT underestimated motor control in pts with severe	it is difficult to confidently accept the positive claims and recommendations pertaining to the AMAT from this review alone. Could not speak to sensitivity or MCID of AMAT. Researchers did not document search strategy or results of the review. It would be difficult to replicate this review and its comprehensiveness cannot be determined.

					impairment.	
Rowland & Gustafsson (2008).	Review the literature of UE assessments used for stroke pts, analyze assessments to determine clinical utility, and rate each assessments psychometrics data.	Descriptive, systematic review; Pyramid level: D1, AOTA level: I	*N*=7 assessments including WMFT, AMAT, ARAT, MAL, UL-MAS, ABILHAND, & CAHAI. Criteria: UE ability assessment, ≥1 test item tests ADL, psychometric data included pts with stroke, measures that use quantitative data, assessments administered by occupational therapists.	Criteria for measure analysis: Purpose of measurement, clinical utility, measurement construction, standardization, reliability, and validity. Assessments analyzed using modified criteria of Outcome Measures Rating Forms and Guidelines.	Psychometrics of all measures rated as adequate to excellent. Reliability for AMAT based on studies included in review: adequate. Validity for AMAT rated excellent. AMAT detected difference in change with passing of 1 vs. 2 weeks. Floor and ceiling effects for clients with severely impaired or near-normal UE function. ARAT able to detect clinically relevant change at 5.7 points. ARAT responsiveness indices: rho=0.66, d=0.52. ARAT responsiveness ratio: 2.03.	Studies that looked at AMAT excluded participants unable to extend their wrist ≥20 degrees and MP joints ≥10 degrees. Study lacked thorough discussion of the search strategy, especially with respect to articles or studies included in this review. Studies could have been selected to focus on strengths and minimize weaknesses of the measures.

Abbreviations Key: ADL - activities of daily living, AMAT - Arm Motor Ability Test, AMAT-9 - 9-item Arm Motor Ability Test, AOTA - American Occupational Therapy Association, ARAT - Action Research Arm Test, BI - Barthel Index, CAHAI – Chedoke Arm and Hand Activity Inventory, CMSA - Chedoke McMaster Stroke Assessment, FA - functional ability, FAT - Frenchay Arm Test, FIM - Functional Independence Measure, FMA - Fugl-Meyer Assessment, IP - interphalangeal, MAL – Motor Activity Log, MC - metacarpal, MI - Motricity Index, MM - manual muscle, MMSE - Mini Mental State Examination, PROM - passive range of motion, pt - patient, QOM - quality of movement, SIS - Stroke Impact Scale, UE - upper extremity, UE-MAS – Upper Limb Motor Assessment Scale, WMFT - Wolf Motor Function Test

The Chedoke Arm and Hand Activity Inventory Quantitative Results

Author, Year	Study Objectives	Study Design/ Level of Evidence	Participants: Sample Size, Description Inclusion and Exclusion Criteria	Interventions & Outcome Measures	Summary of Results	Study Limitations
Barecca, Stratford, Lambert, Masters, & Streiner (2005)	Estimate test-retest reliability and validity of the CAHAI. Test if CAHAI was more sensitive to change in upper limb function than CMSA or ARAT.	Association/ correlation; Pyramid level: D2, AOTA level: IV	*N*=39. Participants had suffered a stroke and receiving therapy in an inpatient or outpatient facility. 19 female, 20 male. *n* = 15 (chronic/ severe, *n* = 24 (acute/mild-moderate stroke). Mean age (acute) 71.4 yrs. Mean age (chronic) 64 yrs. Mean time since onset (acute) 27.3 days. Mean time since onset (chronic) 101.7 days. Exclusion: LMN injury, UE injury pre-stroke, not enough stamina to participate.	Pts were assessed using the CMSA and the CAHAI at baseline and 2 to 6 wks later. The ARAT was used at baseline and follow-up. ICC was determined. Construct and discriminant validity of CAHAI was assessed.	ICC for CAHAI was .98, which shows high interrater reliability. Convergent and discriminant validity of CAHAI was supported. It was found that CAHAI is better at distinguishing change than the CMSA arm-hand sum and the ARAT.	Small sample size. Time of reassessment varied a lot per pt. Therapy setting varied among pts.
Barreca, Stratford, Masters, Lambert & Griffiths (2006)	To determine if the validity of scores were greater for 2 versions of the CAHAI or for the ARAT. To determine if	Association/ correlation; Pyramid level: D2, AOTA level: IV	*N*=105. Participants were pts at 1 of 4 rehab facilities. 51 female, 54 male. Inclusion: stroke pt. Exclusion: LMN injury, UE injury pre-stroke, not enough	CAHAI-13 was administered to participants upon initial eval and discharge from OT or PT. A research therapist also administered the	Both versions of CAHAI had identical levels of cross-sectional validity. Receiver operating characteristic curve areas significantly greater for CAHAI	Pts were not actually administered the CAHAI-9, yet results were gathered as if they did. Pts received therapy in different

	validity of the scores on the CAHAI-13 were greater than the CAHAI-9.		stamina, combined arm and hand CMSA score >11.	CAHAI-13, as well as the ARAT within 36 hrs of the therapist. CAHAI-9 scores were abstracted from CAHAI-13 scores because the 9 tasks from the CAHAI-9 assessment are included in the CAHAI-13 assessment.	versions compared to ARAT, showing greater sensitivity to change.	settings, thus creating another variable, though it could also improve generalization.
Barreca, Stratford, Masters, Lambert, Griffiths, McBay (2006)	To determine reliability and validity of 3 shortened versions of the CAHAI.	Association/ correlation; Pyramid level: D2, AOTA level: IV	*N*= 39; *n* (acute stroke)= 24, *n* (chronic stroke)= 15. Exclusion: LMN injury, UE injury pre-stroke, CMSA score > 11, not enough stamina. Median number of days since stroke 27.3 acute, 101.7 chronic. Median age 71 yrs acute, 64 yrs chronic.	Item reduction of CAHAI-13 was conducted in order to create 3 shortened versions. Treating therapist and research therapist both administered CAHAI-13 upon eval. Research therapist administered ARAT and treating therapist administered CMSA upon eval. Later, after pt received therapy, research therapist administered ARAT and treating therapist administered CMSA and CAHAI-13. Data gathered was used to	Internal consistency: CAHAI-13: .98, CAHAI-9: .98, CAHAI-8: .98, CAHAI-7: .97, thus the test items relate well to each other. Test-retest reliability: CAHAI-13: .98, CAHAI-9: .97, CAHAI-8: .97, CAHAI-7: .96. High test-retest reliability for all. ROC curve: CAHAI-13: 0.95, CAHAI-9: 0.94, CAHAI-8: 0.93 CAHAI-7: 0.97. All very sensitive to change. Clinicians who administered CAHAI recommended	Small sample size. Setting was not mentioned and may have varied among pts.

				assess cross-sectional and longitudinal validity of shortened versions of CAHAI.	the CAHAI-9, as it included important tasks that the shorter version cut out, while still maintaining accurate statistical findings.	
Rowland, Gustafsson, Turpin, Henderson, & Read (2011)	To investigate clinical utility of CAHAI-9 in an acute stroke setting.	Association/ correlation; Pyramid level: D2, AOTA level: IV	*N*= 32. Participants were occupational therapists from 8 hospitals working in acute medical or stroke setting. 31 female, 1 male. Yrs of neuro experience: <1 yr: 13, 1-5 yrs: 10, 5+ yrs: 9.	Participants completed 100 CAHAI-9 assessments on 92 pts with stroke over 6 mos. They also completed questionnaires regarding clinical utility of CAHAI-9. Questions included: mean administration time, # of pts assessed by bedside, and mean scores.	Test item 9 (cutting putty) presented the greatest difficulties- 53% said overall, it was too difficult for pts to perform. Test item 2 (dialing 911) was not applicable to Australian pts. 75% of participants planned to use CAHAI-9 in the future. Mean administration time for CAHAI: 16.2 minutes.	Clincians were relatively inexpereinced in neuro, only 25% had main case load of acute stroke.

Abbreviations Key: ARAT- Action Research Arm Test, CAHAI- Chedoke Arm and Hand Activity Inventory, CMSA- Chedoke-McMaster Stroke Assessment, eval - evaluation, hr - hour, ICC – intraclass correlation coefficient, LMN – lower motor neuron, pt - patient, UE – upper extremity

The Chedoke Arm and Hand Activity Inventory Qualitative Results

Author, Year	Study Objectives	Study Design/ Level of Evidence	Participants: Sample Size, Description Inclusion and Exclusion Criteria	Methods for enhancing rigor	Themes and Results	Limitations
Gustafsson, Turpin, Dorman (2010)	To explore the clinical utility of the CAHAI when used by occupational therapists in stroke rehabilitation.	Qualitative; Pyramid level: Q3 rigor-(b, d), AOTA level: V	*N*=13. 10 female, 3 male. Participants were occupational therapists working in stroke rehab settings. 8 participants had been trained in administration of the CAHAI (group A), 5 had not (Group B). Participants had been working for 7.5 ± 8.5 yrs.	Participants took part in 25 min focus group. They watched a video of a client being assessed using the CAHAI. They then scored the client. Discussion questions about utility of CAHAI followed. Thematic coding categories were determined from the discussion portion by 3 independent raters.	Themes: "Instructions ambiguous and scoring unclear", "how we use it", "whole task versus motor components", "knowing when to use it", "detecting other impairments", "changing the way clients do tasks" and "how they analyze tasks". Some therapists use the CAHAI as an intervention tool. Some therapists used components of the CAHAI to assess factors other than UE impairments.	Clinicians did actually not administer the CAHAI. Study was only 25 minutes in length, thus some themes may be missing.

Abbreviations Key: CAHAI- Chedoke Arm and Hand Activity Inventory, min – minute, pt - patient, UE – upper extremity, yr - year.

The Motor Evaluation Scale for Upper Extremity in Stroke Patients Results

Author, Year	Study Objectives	Study Design/ Level of Evidence	Participants: Sample Size, Description Inclusion and Exclusion Criteria	Interventions & Outcome Measures	Summary of Results	Study Limitations
Van de Winckel, Feys, Messerli, Baronti, Lehmann, Hemelrijk, Pante, Perfetti, & Weerdt (2006)	Explored the validity and unidimensionality of the MESUPES using Rasch analysis as well as interrater reliability.	Outcomes pre-post study; Pyramid level: O4, AOTA level: IV	*N*= 396 in the Rasch study, 158 female, 238 male, and from that sample, *N* = 5, 25 female, 31 male in the reliability study. Mean age 63-65 yrs. Average time since stroke 7.18 +/-16.36 mos. Participants were pts at one of 12 facilities in either Germany, Belgium or Switzerland. Inclusion: 1) No other neurological conditions, orthopedic, or rheumatic impairments, 2) No contractures of the UE, 3) No neglect or apraxia, 4) Able to understand and follow instructions.	All pts were tested by 1 of 7 evaluators who each received one hour of MESUPES training. Rasch Analysis performed on arm function and hand function components separately. 56 of those pts were evaluated by two assessors to determine interrater reliability; second assessment was administered 24 hrs after the first. Interrater reliability was investigated using weighted percentage agreement, ICC, and kappa values.	MESUPES measures the quality of movement. Weighted percentage agreement (>85.7%) and Kappa values were high (0.62-0.79) suggesting high reliability. Initial analysis determined MESUPES is a multidimensional measure. All response categories demonstrated logical order of increasing difficulty. Three hand items and two arm items did not fit the Rasch model. After removing those items, both subscales fit the model suggesting good internal construct validity and unidimensionality.	Average age of pts in this study is young compared to the general population of stroke pts. This study did not include comparisons to other scales in the analysis. Study was performed in Europe.
Johansson	Determine inter-	Descriptive	*N*= 42. 15 female,	An updated version of	Linear-weighted k	Moderate sample

& Häger (2012)	rater reliability, estimate the MDC, and investigate the concurrent validity of the MESUPES.	correlational; Pyramid level: D2, AOTA level: IV	27 male. Mean age 56 yrs. Most in chronic phase. Median time post-stroke 7 mos, mild-to-severe deficits. Inclusion: 1) No other neurological, orthopedic, or rheumatic conditions, 2) No UE contractures, 3) No neglect or apraxia, 4) Ability to understand and follow instructions.	the MESUPES was used, breaking the assessment into an arm scale and a hand scale. All pts tested on the MESUPES by two (out of four) random, trained physiotherapists. Second test administered within 48 hrs of the first. Inter-rater reliability determined by percentage of maximum one-score difference, linear-weighted k analysis, and percentage of agreement. Spearman's rho calculated between MESUPES and UE MAS to determine concurrent validity. MDC calculated using SEM and assessment data.	values (0.63-0.96) indicating good to very good inter-rater reliability. K-value of >0.86 on 10 out of 17 items. Score differences on items was $\leq$2. Very high correlation based on ICC of >0.98 for subscores and total score. Sufficient absolute reliability determined by SEM of 2.68. Correlation score of r = 0.87 for MESUPES and the UE M MAS indicate good concurrent validity. Score change of 5, 7, and 8 is required to achieve 80%, 90%, and 95% confidence. MDC% for MESUPES is sufficiently low (18%).	size of chronic phase stroke survivors means generalizability is limited. This study did not consider intra-rater reliability or sensitivity of the MESUPES. Study was completed in Europe. Raters had no prior experience with this measure. Raters had two hours of training.

Abbreviations Key: AOTA – American Occupational Therapy Association, hrs – hours, ICC – intra-class correlation, K-value – kappa value, MDC – minimal detectable change, MESUPES – Motor Evaluation Scale of Upper Extremity in Stroke Patients, M MAS- Modified Motor Assessment Scale, pt = patient, SEM – standard error of the mean, UE - upper extremity

The Streamlined Wolf Motor Function Test Results

Author, Year	Study Objectives	Study Design/ Level of Evidence	Participants: Sample Size, Description	Interventions & Outcome Measures	Summary of Results	Study Limitations

			Inclusion and Exclusion Criteria			
Bogard, Wolf, Zhang, Thompson, Morris, Nichols-Larsen (2009)	Determine if it is possible to streamline the WMFT by defining only those items that showed greatest improvement relative to the entire test.	Descriptive; Pyramid level: D2, AOTA level: IV; data drawn from larger RCT (E2, I)	*N* = 169; *n*(immediate group) = 96, *n(*delayed group) = 73. All participants in EXCITE trial from 6 university programs. Inclusion: 1st time stroke, thumb and 2 digits ≥ 10°, transfer independently, stand for 2 min; no cognitive impairments or unstable medical status, only participants who completed all tx.	Participants randomized into immediate tx group or delayed tx (I yr later) group; tx involved safety mitt on less affected limb for 90% of waking hours for 14 consecutive days and ATP and RTP progressing to 6 hrs/day for 10 days total; WMFT given by trained therapists at baseline, post-tx, and every 4 mo for 2 yrs, videotaped for rating by independent therapists; covariates included gender, functional level, concordance.	Log mean scores for change scores calculated, covariates controlled for; immediate group showed significant ($p < 0.05$) relationships with 6 timed items: hand to table (front), hand to box (front), reach and retrieve, lift can, lift pencil, fold towel; delayed group showed significant ($p < 0.05$) relationships with 6 timed items; extend elbow weight hand to box (front), lift can, lift pencil, turnkey in lock, fold towel. The authors conclude that the two streamlined versions of the WMFT are viable for assessing function and tracking change.	Little demographic information provided, mild to moderate stroke severity, delayed group received other tx during yr before CIMT, no reliability or validity done on proposed streamlined version, EXCITE inclusion criteria not met by most stroke survivors; streamlined version specific to stroke recovery stages.
Chen, Wu, Lin, Chen, Chen, & Chen (2012)	Examined psychometrics of subacute and chronic versions of the S-WMFT.	Outcomes, pre-post (part of an RCT E2, I); Pyramid level: O4, AOTA level: III	*n*(chronic phase stroke)= 97. 33 female, 64 male. Median age 56.6 yrs, median time post-stroke 20	Three trained occupational therapists administered WMFT before and after participation in	Unidimensionality determined. No significant DIF by demographics found in either version. Subacute stroke version did not measure UE motor function	Sample size is moderate and limited to mild and moderate stoke severity, which limits

			mo. *n*(subacute stroke)= 75. 17 female, 58 male. Median age 55.6 yrs, median time post-stroke 7 mo. Inclusion: First-time stroke with onset 3-9 mo. prior (subacute) or ≥12 mo. prior (chronic), Ability to understand and respond to questions, MMSE score ≥22, Brunnstrom stage ≥ 3 for proximal affected UE, score ≤ 2 on Modified Ashworth Scale.	rehabilitation programs. Outcomes included test structure, targeting, reliability, and item difficulty hierarchy. Construct validity, unidimensionality, and correlations between items were also investigated through data (and Rasch) analysis. Researchers tested DIF to determine how demographics influence test outcomes.	as well as chronic version. Item correlation coefficient range: 0.37-0.74 meaning that the items measure similar UE motor components. Some redundancy of items was found in the subacute version and the authors recommended removal of 1of the 2 highly correlated items (move hand to box and move hand to table). Both versions are useful measures for pts with mild to moderate impairment post-stroke.	generalizability to facilities that also treat more severe stroke survivors. Did not include participants with severe impairments, also limiting the external validity.
Chen, Wu, Lin, Jang, Lin, Cheng, Chung, & Yan (2014)	Can the SWMFT-C and SWMFT-S be used in for pts >12 mo and 3-9 mo post-stroke, do the two tests cover a continuum of recovery?	Association/ correlation; Pyramid level: D2, AOTA level: IV	*N* = 351, drawn from ongoing studies at 6 hospitals. Inclusion: Brunnstrom ≥ 2 proximal and distal, MMSE ≥ 18, first time stroke, no major medical problems (aphasia, visual deficits, poor	Participants received one of following txs for 90 min/day for 3 or 4 wks: conventional rehabilitation, BAT, dCIT, robot-assisted therapy, mirror therapy; 17-item WMFT administered by trained independent therapists; 4	Unidimensionality: PCA of residuals showed that the Rasch model explained 72.2% of the varience, indicating unidimensionality; Rasch analysis also found the item hierarchy of the SWMFT was similar to the WMFT (items were ordered from least to most difficult); average difficulty of SWMFT items covered full	Only assessed FAS of SWMFT, some motor return present, no severe cognitive impairments, limited to mild to moderate stroke severity, took place in Taiwan.

			physical condition), split into 3 groups: acute (mean age 53.23 ±12.83 yrs), 9-12 mo (53.67 ± 11.46), chronic (54.85 ± 11.23).	overlapping items from two SWMFT tests and 4 unique items combined into 8 items reviewed in this study.	range of UE function (from -1.49 to 1.59 logits), overall range of SWMFT items -8.38 to 7.80 logits; both the SWMFT-S and the SWMFT-C can be used accurately with pts 9-12 mo. post-stroke.	
Wu, Fu, Lin, Feng, Hsieh, Yu, Lin, Hsieh, & Ota (2011)	Determine responsiveness, concurrent validity, and predictive validity of the S-WMFT as compared to the WMFT.	Descriptive; Pyramid level: D2, AOTA level: IV	*N* = 64, 54 participants from previous study by authors, 10 additional recruited from rehab departments at 3 hospitals. Inclusion: first time stroke 3-10 mo. prior, Brunnstrom ≥ III for proximal affected limb, MMSE > 24, MAS ≤ 2, no other major medical problems (apahsia, low vision, poor overall health). Mean age 53.01 ± 12.75 yrs, mean time since stroke 7.88 ± 1.69 mo.	Randomly assigned to 1 of 3 interventions for 2 hr/day for 3 wks: dCIT, BAT, or conventional therapy; evaluated at baseline and immediately after tx by 3 trained OTs; responsiveness found using Wilcoxon signed rank test, validity by correlation with the FMA and SIS-Hand (Spearman rank correlation coefficient), predictive validity by tracking change from baseline to post-test.	Responsiveness: both tests small to moderate (S-WMFT effect size $d = 0.41$, WMFT effect size $d = 0.52$), no statistically significant difference between the two tests Concurrent validity: S-WMFT had good correlation with FMA ($\rho =$ 0.69 at baseline, 0.58 at post-test) and good with SIS-Hand ($\rho = 0.59$ at baseline, 0.57 at post-test); WMFT had good correlation with FMA ($\rho =$ 0.64 at baseline, 0.64 at post-test) and with SIS-Hand ($\rho = 0.51$ at baseline, 0.54 at post-test) Predictive validity: both test correlated significantly with the FMA and SIS-Hand ($\rho \geq$ 0.56, $p < 0.01$), very little difference between the two tests. Authors conclude that	No cognitive impairments, high level of arm function present, limited range of time since stroke; no reliability data included, sub-acute stroke only.

					both versions of the S-WMFT are acceptable measures for detecting change in movement ability and predicting return of motor function. Compared to the original test, the streamlined tests have similar sensitivity and concurrent validity, and better predictive validity and clinical utility.	

Abbreviations Key: ATP – adaptive task practice, BAT – bilateral arm training, dCIT – distributed constraint-induced therapy, DIF – differential item functioning, FMA – Fugl-Meyer Assessment, hr - hour, MAS – Modified Ashworth Scale, min - minute, MMSE – Mini-Mental State Examination, mo. – month, OT – occupational therapist, PCA – principal component analysis, pt - patient, RCT – randomized controlled trial, rehab – rehabilitation, RTP – repetitive task practice, SIS-Hand – hand section of the Stroke Impact Scale, S-WMFT – Streamlined Wolf Motor Function Test, SWMFT-FAS – Streamlined Wolf Motor Function Test Functional Ability Scale, SWMFT-C – Streamlined Wolf Motor Function Test Chronic, SWMFT-S – Streamlined Wolf Motor Function Test Sub-acute, tx - treatment, UE – upper extremity, wk - week, WMFT – Wolf Motor Function Test, yr – year

Stroke Descriptive Review Results

Author(s), Year	Study Objectives	Study Design/ Level of Evidence	Assessments or Screens Being Compared	Psychometrics	Population/ Setting	Summary of Results	Limitations
Lang, Bland, Bailey, Schaefer, & Birkenmeier, 2013.	To provide an approach for assessing the UE post-stroke.	Descriptive review; Pyramid level: D3, AOTA level: V	ARAT, BBT, CAHAI, JTT, NHPT, WMFT, MAL, SIS.	Interrater reliability: ARAT- .98, BBT- .99, CAHAI- .98, JTT- 0.82 to 1.00, NHPT- 0.68 to 0.99, WMFT- 0.85 to 0.97. Test-retest reliability: ARAT- .98, BBT- .96, CAHAI- 0.96 to 0.97, NHPT- 0.68 to 0.99, WMFT- 0.94 to 0.99, MAL- 0.79 to 0.82, SIS- 0.70 to 0.92.	All of these assessments are for use on pts with stroke and can be performed in any rehab setting.	The assessments are all strong and all measure UE function. All of them are recommended for clinical use. 3 questions help choose appropriate assessment: Is equipment available? Is training needed? How long to administer?	Nothing is mentioned specifically about the studies that were done in order to glean psychometric data for the assessments, thus limitations are unknown. Inclusion and exclusion criteria were not discussed, and there was limited data provided on the included studies.
Poole & Whitney (2001)	Review assessments of motor function that are appropriate for post-stroke pts.	Systematic Review; Pyramid level: D1, AOTA level: V	FMA, MAS, RMA, CMSA, MCA, MI, STREAM, FAT, ARAT, FT, and AMAT.	Interrater reliability: FMA-.96, MAS- .95 to .99, RMA- t = .80, CMSA- .97, MI- spearman's rho = .88, STREAM-GCC= .96 to.99, FAT- 100% agreement, ARAT- r = .98 and .99, FT- spearman's rho = .97, AMAT- spearman's rho =	All of these assessments are for use on pts with stroke and can be performed in any rehab setting.	Test-retest reliability is high for the AMAT compared to others such as the FAT and CMSA, and ARAT. Intra-rater reliability was at least adequate for all 11 assessments. FMA, RMA, CMSA, and FT	Lacks discussion of how articles and/or studies were selected to be included in this review including the samples, interventions, and other influential aspects of studies that could have influenced the psychometric data.

				.69 to.99. Intrarater: FMA- r = .98, MAS- r = .98, CMSA- .98 total score, ARAT- r = .99.		take long to administer, not advised. FMA, RMA, CMSA, and AMAT more sensitive to change than MAS, MCA, MI, STREAM, FAT, ARAT, and FT. It is recommended that an assessment of motor activity ability be used concomitant with functional participation assessments such as the Barthel and FIM.	Lacks discussion of the process of determining which assessments fit the inclusion criteria as well as a lack of discussion in terms of the quality of each study.

Abbreviations Key: AMAT- Arm Motor Ability Test, ARAT- Action Arm Research Test, BBT- Box and Block Test, CAHAI- Chedoke Arm and Hand Activity Inventory, CMSA- Chedoke-McMaster Stroke Assessment, FIM- Functional Independence Measure, FMA- Fugl-Meyer Assessment, FT- Functional Test for the Hemiparetic Extremity, GCC- generalizability correlation coefficient, JTT- Jebson-Taylor Hand Function Test, MAL- Motor Activity Log, MAS- Motor Assessment Scale, MCA- Motor Club Assessment, MI- Motricity Index, NHPT- Nine-Hole Peg Test, RMA- Rivermead Motor Assessment, SIS- Stroke Impact Scale, STREAM- Stroke Rehabilitation Assessment of Movement, WMFT- Wolf Motor Function Test.

Stroke Systematic Review Results

Author, Year	Study Objectives	Study Design/ Level of Evidence	Number of Papers Included, Inclusion and Exclusion Criteria	Interventions & Outcome Measures	Summary of Results	Study Limitations
Murphy, Resteghini, Feys, & Lamers (2015)	To discover Psychometric properties and clinical utility of UE outcome measures for stroke pts via systematic review.	Systematic review (of systematic reviews); Pyramid level: D1, AOTA level: I	Systematic search included systematic reviews from 2004 to February 2014. 13 systematic reviews were reviewed after searching 7 databases. Search criteria: humans, English language, adult: 19+ yrs, field: title and abstract. Inclusion: published in peer-reviewed journals, participants had stroke, reported a clear objective to identify OM specific for the UE, report psychometric properties of the OM.	In 13 reviews, 53 OMs were addressed. FMA, ARAT, BBT, CAHAI, WMFT, MAS, MI, FAT, CMSA, NHPT, STREAM, kinematic measures and ABILHAND were included. AMSTAR-tool was used to appraise the reviews.	6 OMs had a high level of measurement quality and clinical utility: FMA, ARAT, BBT, CAHAI, WMFT, and ABILHAND. Authors recommend use of these OMs for research and clinical practice over all other OMs reviewed.	Conclusions were based on conclusions made by the authors of the original systematic reviews. Overlap of articles used in systematic reviews. Technology assisted OMs were excluded. Limited info was found from primary studies that OMs were extracted from.

Abbreviations Key: AMSTAR- Assessment of Multiple Systematic Reviews, MMAC- Modified Motor Assessment Chart, MAS- Motor Assessment Scale, MCA- Motor Club Assessment, MI- Motoricity Index, NHPT- Nine Hole Peg Test, OM – outcome measure, RMA- Rivermead Motor Assessment, BBT- Box and Block Test, STREAM- Stroke Rehabilitation Assessment Movement, FAT- Frenchay Arm Test, CMSA- Chedoke-McMaster Stroke Assessment.

Parkinson's Disease Assessments

Parkinson's Disease Systematic Review Results

Author, Year	Study Objectives	Study Design/ Level of Evidence	Participants: Sample Size, Description Inclusion and Exclusion Criteria	Interventions & Outcome Measures	Summary of Results	Study Limitations
Proud, Miller, Bilney, Balachandran, McGinley, & Morris (2015)	1) Identify assessments appropriate for pts with PD that measure UE impairment, activity limitation, or restriction in participation. 2) Determine psychometrics and clinical utility of the measures for pts with PD. 3) Use ICF framework to compare assessments.	Descriptive, systematic review; Pyramid level: D1, AOTA level: I	*N*=18 assessments (7 PD and movement disorder specific tests, and 11 for other conditions) were identified. *N*=10 studies that identified assessment psychometrics. Inclusion: knows English, Dx of idiopathic PD, Evaluated tx or disease progression, quantified UE impairment, activity limitation, and/or participation restriction, Described test protocol that could be replicated, provide sufficient detail of assessment(s), no kinematic analyses.	Assessments were identified by searching 9 databases. Studies were identified by searching 5 databases. Checklist developed by Terwee et al and COSMIN checklist were both used to determine quality of assessments and studies.	Most assessments could be finished within 10 min. and cost <$100. Only the UPRDS requires training. Inter-rater and intra-rater reliability of certain UPDRS UE items is moderate. Fair support for test/re-test, reliability of PPT, MPS, and CAPIT. No high-quality studies investigated responsiveness or validity. Evidence of reliability for UPDRS and CAPIT, but little support for validity and responsiveness. Little evidence of test-retest reliability for PPT. Moderately strong between disease severity and PPT scores.	Most studies included <25 participants (range: 12-411) with duration of the disease >5 yrs, meaning that external validity is relatively weak and might only apply to pts in relatively more progressed stages of PD. Study may lack comprehensiveness since many timed UE assessments excluded due to lack of details or description in studies, or equipment not available commercially. Limited responsiveness and validity studies could limit the clinical utility of the measures addressed in this review.

Abbreviation Key: AOTA – American Occupational Therapy Association, CAPIT – Core Assessment Program for Intracerebral Transplantations, COSMIN – consensus-based standards for the selection of health status measurement instruments, Dx – diagnosis, ICF – International Classification of Functioning, Disability, and Health, min. – minute(s), MPS – Motor Performance Series, PD – Parkinson's disease, pts – patients, PPT – Purdue Pegboard Test, tx – treatment, UE – upper extremity, UPDRS – Unified Parkinson's Disease Rating Scale

The Rush Dyskinesia Rating Scale Review Results

Author, Year	Study Objectives	Study Design/ Level of Evidence	Participants: Sample Size, Description Inclusion and Exclusion Criteria	Interventions & Outcome Measures	Summary of Results	Study Limitations
Colosimo, Martinez-Martin, Fabbrini, Hauser, Merello, Miyasaki, Poewe, Sampaio, Rascol, Stebbins, Schrag, & Goetz (2010).	Compare psychometrics of assessments that measure drug-induced dyskinesia in pts with PD to make recommendations regarding clinical utility.	Descriptive, systematic review; Pyramid level: D1, AOTA level: I	*N*=8 assessments; identified by systematic literature search and consensus among task force members following the same methods as the task forces that critiqued rating scales for mental health in PD.	Psychometric data was summarized and clinical utility was addressed for each measure. Measures included AIMS, UPDRS, Obeso (CAPIT), RDRS, CDRS, Lang-Fahn, PDYS-26, and UDysRS. Researchers assigned each assessment as either "recommended", "suggested", or "listed" depending on which of the following criteria are met: if it has been studied clinimetrically and found to have sound psychometric data, if there were studies authored by individuals not involved in the test's development, and if it had been applied to	Only RDRS and AIMS received a rating of "recommended". Six measures were rated as "suggested" and no measures were rated as "listed". A caveat to the rating system: some scales rated as "suggested" have very strong psychometrics, however, they were too new to position them in "recommended" status. High inter- and intra-rater reliability for RDRS with good clinical utility (e.g., time to administer). May be difficult to tease out specific dyskinesia with RDRS and does not assess disability in the context of wide variety of functional	Study lacked thorough discussion of the search strategy, especially with respect to articles or studies included in this review. Select studies could have been selected to focus on strengths and minimize weaknesses of the measures. Limited search using only 2 databases to find articles. Very few specific psychometric data listed, especially for newer measures (e.g., No discussion of MCID or sensitivity). This

				pts with PD.	activities.	can make comparing measures difficult.
Goetz, Stebbins, Chung, Hauser, Miyasaki, Nicholas, Poewe, Seppi, Rascol, Stacy, Nutt, Tanner, Urkowitz, Jaglin, & Ge (2013).	Objectives of this RCT included comparing the sensitivity of measures that capture data on dyskinesia secondary to PD, determine effect sizes, and to explore placebo-related immunities.	Double-blind randomized controlled trial; Pyramid level: E2, AOTA level: I	N=60 pts statistical power was achieved at 50). Inclusion criteria: Dx of PD, Age: 30-90 yrs, permanent dyskinesia (CGI_{-s} score $\geq$3), documented normal creatinine level, no tx with amantadine >3 mos., stable doses of all antiparkinsonian medications for $\geq$4 wks, caregiver willing to participate, no brain surgery, urinary retention, glaucoma, depression, dementia or hallucinations.	Pts with PD were randomly assigned to either amantadine or placebo treatment groups. Pts participated for 8 wks. One pre- and 2 post-evaluations were administered using UDysRS, LF, PDD-26, RDRS, AIMS, dyskinesia items from MDS-UPDRS, and CGI-C and CGI-S. Order in which measures were administered was randomized for each visit. Sensitivity to change was analyzed using a repeated-measures ANOVA.	Only four measures (CGI-C, LD, PDD-26, & UDysRS) were sensitive to changes in dyskinesia between tx and placebo groups; not RDRS. Small effect size for RDRS (0.003) compared to UDysRS (0.138). UDysRS was most sensitive; could detect a tx change 80% of SD. UDysRS combines aspects of RDRS and AIMS, neither of those latter measures could detect tx changes independently. Intraclass correlations for UDysRS were higher when pts were rated face-to-face as opposed to video camera. No assessments were immune to placebo effects (attributed to biochemistry of PD and trial design).	While it was a double-blind study, measures were known to raters as they were administered which could have resulted in biased scoring. Placebo group had significantly higher CGI-S scores at baseline. The number of scales administered during each visit (8) could have caused fatigue.

Abbreviations Key: AIMS – Abnormal Involuntary Movement Scale, CDRA – Clinical Dyskinesia Rating Scale, CGI-C – Clinical Global Impression (change), CGI-S – Clinical Global Impression (severity), MCID – Minimal Clinically Important Difference, Dx – diagnosis, LF – Lang-Fahn Activities of Daily Living Dyskinesia Rating Scale, MDS-UPDRS – Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale, PD – Parkinson's disease, PDD-26 – Parkinson's Disease Dyskinesia Scale, PDYS-26 – Parkinson Disease Dyskinesia Scale, pts – patients, RCT – Randomized Controlled Trial, RDRS – Rush Dyskinesia Rating Scale, SD – standard deviation, tx – treatment, UPDRS – Unified Parkinson's Disease Rating Scale, UDysRS – Unified Dyskinesia Rating Scale, wks – weeks, yrs – years

Summary of Findings

Overall, 33 studies on the psychometric properties of UE assessments for stroke and three for PD were identified for inclusion in this appraisal. The studies identified were conducted in limited settings and with limited populations. Many articles did not include patients with cognitive impairment, which limits the generalizability of the results to clinical populations. Many of the stroke studies also required a certain level of motor return to participate, which resulted in less information on suitability of the measure for more severely impaired clients. The studies on PD also lacked diversity in the study populations in terms of severity, and little fluctuation or dyskinesia. Within the studies, many of the assessments were administered by physical therapists. It is unclear whether there would be any differences related to being administered by occupational therapists

Table 7

Summary of Key Finding for Stroke Assessments

	ARAT	AMAT
Brief description	As outlined by Yozbatiron, Der-Yeghiaian, and Cramer (2008), the test consists of 19 tasks split into 4 subscales (grasp, grip, pinch, gross motor) and scored on a 4 point ordinal scale based on task performance. While other authors have suggested changes to the test, we recommend following the structure and format outlined by the aforementioned authors for consistency in administration.	Measures qualitative and quantitative aspects of 13 ADL that each involves 1-3 movements or components. Movements and components are performed continuously; therefore, this assessment does not interfere with natural movements. Non-affected UE can be used to stabilize. Each task has a 1-2 minute time limit. Patients are rated based on ability to perform tasks and quality of movement. (O'Dell et al., 2011)
Actual component tested/aim of measure	Recovery of UE function and motor status after stroke.	Functional UE ability and quality of movement.
Equipment Required	1. 2 metal tubes (different sizes) 2. 4 blocks (different sizes) 3. Cricket ball 4. Marble 5. Sharpening stone 6. Tobacco can/lid 7. Steel washer 8. Plastic pitcher 9. 3 ball bearings (different sizes) 10. Smoothing iron 11. Plastic cups	1. Silverware & plate 2. Play-doh 3. Foam sandwich 4. Mug 5. Comb 6. 2 styles of shirts 7. Towel 8. Jar 9. Door 10. Light switch 11. Kidney beans 12. Telephone 13. Shoe/shoelace
Test items	Grasp subscale: 1. Block, 10 cm^3	1. Cut "Meat" 2. Foam "sandwich"

	2. Block, 2.5 cm^3 3. Block, 5 cm^3 4. Block, 7.5 cm^3 5. Cricket ball 6. Sharpening stone Grip subscale: 7. Pour water from one glass to another 8. Displace 2.25-cm alloy tube from one side of table to the other 9. Displace 1-cm alloy tube from one side of the table to the other 10. Put washer over bolt Pinch subscale: 11. Ball bearing, held between ring finger and thumb 12. Marble, held between index finger and thumb 13. Ball bearing, held between middle finger and thumb 14. Ball bearing, held between index finger and thumb 15. Marble, held between ring finger and thumb 16. Marble, held between middle finger and thumb Gross movement subscale: 17. Hand to behind the head 18. Hand to top of head 19. Hand to mouth	3. Eat with spoon 4. Drink from mug 5. Comb hair 6. Open jar 7. Tie Shoelace 8. Use telephone 9. Wipe up spilled water 10. Put on cardigan (jacket-style) sweater 11. Put on T-shirt 12. Prop on extended arm 13. Light switch/door (Kopp et al, 1997)
Scoring	0 – Subject unable to complete any part of the task 1 – Subject partially completes task 2 – Task completed with great difficulty or takes abnormal length of time 3 – Task performed normally	**Functional Ability Scale:** 1. Does not attempt with involved arm. 2. Involved arm does not participate functionally; however, attempt is made to use the arm. In unilateral tasks that uninvolved extremity may be used to move the involved extremity. 3. Does, but requires assistance of uninvolved extremity for minor readjustments or change of position, or requires more than two attempts to complete, or accomplishes very slowly. In bilateral tasks the involved extremity may serve only as a helper or stabilizer. 4. Does, but movement is influenced to some degree by synergy or is performed slowly and/or with effort.

		5. Does, movement is close to normal*, but slightly slower; may lack precision, fine coordination or fluidity. 6. Does, movement appears to be normal.* **Quality of Movement:** 1. Partial range movement accomplished, but: 2. Movement dominated by synergy, or there is gross incoordination between limb segments, or extremity nonfunctional for weight bearing activities. 3. Movement accomplished, but: a. Is influenced by synergy, or is accompanied by excessive compensatory movements of trunk, head, or contralateral upper extremity, or lacks either proximal control or fine motor ability, or movement performed very slowly, or minimally able to perform weight bearing activities. 4. Some isolated movement, but: a. Influenced to some degree by synergy, or movement with little influence of synergy but performed slowly, or moderate incoordination and lack of accuracy, or weight bearing activities are performed with difficulty, or primitive grasp patterns are present. 5. Movement close to normal*, but: a. Slightly slower, or lacks precision, fluidity or precise coordination of movement, or able to perform weight bearing activities but with some hesitancy or mild difficulty. 6. Normal movement*: a. Fluid and coordinated activity, speed of movement appears within normal limits. (*) For the determination of normal the uninvolved limb can be utilized as an available index for comparison, with premorbid limb dominance taken into consideration.ent Scale:
Clinical Utility	Cost: Assessment is free, kit components must be purchased.	Cost: $25; Time: 30-40 min (31-60 min test); Time to administer is less with

	Time: 5-15 minutes (10-20 to set up) Training: None, recommend following the standardized administration outlined by Yozbatiron (Yozbatiron, Der-Yeghiaian, & Cramer, 2008).	items removed. Training: Reading an article/manual.
Reliability	Interrater reliability: ICC 0.92-0.9986; percentage agreement 94% (Nijland, et al., 2010; Yozbatiron, Der-Yeghiain, & Cramer, 2008) Intrarater reliability: ICC 0.97- 0.99; percentage agreement 89-91% (Nijland, et al., 2010; Yozbatiron, Der-Yeghiain, & Cramer, 2008; Nordin, Murphy, & Danielsson, 2014)	Interrater reliability: 0.95, 0.68-1.0, 0.69-0.99 (Kopp, Kunkel, Herta Flor, Platz, Mauritz, Gresser, McCulloch, & Taub, 1997; O'Dell, Kim, Finnen, & Polistena, 2011) Test-retest reliability: 0.93-0.99 (Kopp, Kunkel, Herta Flor, Platz, Mauritz, Gresser, McCulloch, & Taub, 1997) ICC: 0.94-0.97 (O'Dell, Kim, Finnen, & Polistena, 2011)
Validity	Predictive validity: 0.45 (Chen, Lin, Wu, & Chen, 2012) Construct validity: Disordered, moderate with age, minimally with stroke severity, moderate with strength, poor with light touch, poor with pain, poor with spasticity (Chen, Lin, Wu, & Chen, 2012; Lang, Wagner, Dromerick, & Edwards, 2006) Concurrent validity: 0.82-0.97 (Hsieh, Hsueh, Chiang, & Lin, 1998; Nijland, et al., 2010) Conflicting results on floor effect: One article found an effect at day 14, one found no floor effect (Lin, et al., 2009; Nijland, et al., 2010)	Concurrent validity with MI is modest but significant; Significant correlations between AMAT-13, FIM self-care scale, and SIS; Correlation coefficients with ARAT, FMA, and WMFT: 0.78-0.79 (O'Dell, Kim, Finnen, & Polistena, 2011; O'Dell, Kim, Rivera, Fieo, Christos, Polistena, Fitzgerald, & Gorga, 2013)
Internal consistency	Unidimensionality established: Scalability coefficient *H* = 0.79-0.95 (Van der Lee, Roorda, Beckerman, Lankhorst, & Bouter, 2002; Koh, et al., 2006) Cronbach's alpha = 0.985 (Nijland, et al., 2010)	ICC: 0.94- 0.97 (O'Dell, Kim, Finnen, & Polistena, 2011)
Minimal clinical significant difference (minimal detected	MCID: 5.7, 10% (this result from a stronger article) (Van der Lee, et al., 2001) For dominant hand: 12 points, 21% of total, effect size 0.78 (Lang, Edwards, Birkenmeier, & Dromerick, 2008) For non-dominant hand: 17 points, 30% of total, effect size 1.10 (Lang, Edwards, Birkenmeier, & Dromerick,	Differences detected with the passing of one versus two weeks. (Rowland & Gustafsson, 2008).

change)	2008)	
Responsiveness	Lower than FMA (difference in SRM 0.47); effect size 0.49-1.390, Wilcoxon test 4.4-4.5 (Hsieh, et al., 2009)	Responsiveness ranged from not well established to intermediate responsiveness. (O'Dell, Kim, Finnen, & Polistena, 2011; O'Dell, et al., 2013)
Resources	Link to test form: http://www.strokecenter.org/trials/scales/action_research_arm_test.pdf Recommended administration: http://www.ncbi.nlm.nih.gov/pubmed/17704352	Score sheet and guidelines available at: http://www.neuropt.org/docs/edge-documents/stroke-edge-compendium-of-instructions2A86360E9D57.pdf?sfvrsn=2 (search for Arm Motor Ability Test)

	CAHAI	MESUPES	S-WMFT
Brief description	An assessment composed of 13 test items that utilize both UE in functional activities. Each activity is rated on a 7 point scale that measures level of independence via observation of patient performance. Three shortened versions of the assessment exist which include 7, 8 and 9 tasks. (Figueiredo, 2009)	Newer assessment that measures movement performance quality of the hemiparetic upper extremity. Seventeen items assessment with two subtests for arm function: 1) MESUPES-arm (8 items) and, 2) MESUPES-hand (9 items). (McDermott, 2015)	The streamlined version (Bogard et al., 2009) consists of 6 function based tasks ranged from less to most difficult, scored on a 6 point ordinal scale based on bilateral use and quality of movement. Tasks are also timed. There is a version of the test for sub-acute stroke and a version for chronic stroke.
Actual component tested/aim of measure	Recovery of UE function after stroke. (Figueiredo, 2009)	Quality of movement performance post-stroke. (McDermott, 2015)	Characterizes motor status of patients with mild to moderate chronic stroke and TBI
Equipment Required	For CAHAI-9 version: 1. Dycem 2. 200g jar of coffee 3. Push button telephone 4. 12”/30cm ruler 5. 8.5” x 11” paper 6. Pencil 7. 2.3L plastic pitcher with lid 8. 250 ml plastic cup 9. Wash cloth 10. Wash basin (24.5 cm. in diameter, height 8 cm.) 11. Pull on vest with 5 buttons (one side male & one side female) 12. Bath towel (65cm X 100cm) 13. 75 ml toothpastes with screw lid, >50% full 14. Toothbrush	1. Mat 2. Desk 3. Chair (positioned so patient is sitting with hips and knees in 90 degrees flexion). 4. Plastic or wooden block (marked with 1cm and 2cm to measure ROM during hand tasks) 5. Large plastic bottle (cylinder with a diameter of 6cm, such as a 20fl oz soda or water bottle) 6. Small plastic bottle (cylinder, diameter 2.5cm, height 8cm) 7. Dice (1.5 x 1.5 cm) (McDermott, 2015)	1. Chair 2. Box (25.4cm tall) 3. Free-weights 4. Pencil 5. Paperclip 6. Can 7. Table (28cm long) 8. Bedside table 9. Cards 10. Checkers 11. Basket 12. Towel 13. Key lock with key 14. Dynamometer Note: this equipment list is for the full version of the test, not the streamlined versions.

	15. Dinner plate (Melamine or heavy plastic, 25 cm. in diameter) 16. Medium resistance putty 17. Knife and fork 18. Built up handles the length of the utensil handle		
Test items	1. Open jar of coffee 2. Call 911 3. Draw a line with a ruler 4. Pour a glass of water 5. Wring out washcloth 6. Do up five buttons 7. Dry back with towel 8. Put toothpaste on toothbrush 9. Cut medium resistance putty (http://www.cahai.ca)	MESUPES arm scale: 1. Hand to stomach 2. Hand back to starting position 3. abduction 0°-90°, arm extended, forearm in neutral position (arm slides on the table) 4. Arm back to starting position 5. Hand from knee (starting position) onto the table 6. Hand(palm) to mouth (elbow remains on the table) 7. Reach for a plastic bottle on the table at arm's length in front of the patient's midline (grasping the bottle is not required) 8. Hand on top of the head (shoulder in abduction) MESUPES hand scale: 1. Pinch grip 2. Wrist extension 3. Opposition of thumb and little finger 4. Selective extension of 3rd finger 5. Starting position with fingers 4 and 5 slightly spread out; spread index and middle finger simultaneously, sliding on the table 6. Selective extension of 5th finger (McDermott, 2015)	**Sub-acute test items:** 1. Hand to table 2. Hand to box 3. Reach and Retrieve 4. Lift can to mouth 5. Lift pencil from table 6. Fold towel **Chronic test items:** 1. Extend elbow with weight 2. Hand to box 3. Lift can to mouth 4. Lift pencil from table 5. Turn key in lock 6. Fold towel

Scoring	These are general scoring guidelines. Each task has specific guides for scoring as well. 7 (complete independence) - All of the tasks are performed safely, without modification, assistive devices or aids, and within reasonable time. 6 (modified independence) - Activity requires any one or more of the following: an assistive device, more than reasonable time, or there are safety (risk) considerations. 5 (supervision) - The client requires no more help than standby, cueing or coaxing, without physical contact. A helper sets up needed items or applies orthoses. 4 (minimal assistance) - With physical contact the client requires no more than touching, and client expends 75% or more of the effort. 3 (moderate assistance) - Weak limb manipulates and stabilizes during the task. The client requires more help than touching, or expends more than 50% of the effort. 2 (maximal assistance) - Weak limb stabilizes during task. The client expends less than 50% of the effort, but at least 25%. 1 (total assistance) - The client	Passive (scores: 0-1) • Patient: Is asked to let the therapist perform the movement with the affected arm. • Therapist: Performs the task slowly to evaluate the adaptation of the tone to the movement. 0 = No adequate adaptation of tone to the movement *(hyper-or hypotonia).* 1 = Adequate adaptation of tone *(normal tone)* to at least part of the movement. Assisted (score: 2) • Patient: Is asked to help perform the movement. • Therapist: Assists the patient as much as needed to perform the movement normally. Feels if and how much the patient actively contributes to the movement in a normal way. 2 = Participation through normal muscle contraction in at least part of the movement. By him/herself (scores: 3-5) • Patient: Performs the movement without help. • Therapist: Watches how far the patient can move in a normal way. 3 = Performs part of the whole movement normally. 4 = Completes the whole movement normally but performs it slowly or with great effort. 5 = Completes the whole movement normally at normal speed. (McDermott, 2015)	0 – Does not attempt with upper extremity (UE) being tested. 1 –UE being tested does not participate functionally; however, attempt is made to use the UE. In unilateral tasks the UE not being tested may be used to move the UE being tested. 2 – Does, but requires assistance of the UE not being tested for minor readjustments or change of position, or requires more than two attempts to complete, or accomplishes very slowly. In bilateral tasks the UE being tested may serve only as a helper. 3 – Does, but movement is influenced to some degree by synergy or is performed slowly or with effort. 4 – Does; movement is close to normal *, but slightly slower; may lack precision, fine coordination or fluidity. 5 – Does; movement appears to be normal *. (*) For the determination of normal, the less-involved UE can be utilized as an available index for comparison, with pre-morbid UE dominance taken into consideration.

	expends less than 25% of the effort. (http://www.cahai.ca)		
Clinical Utility	Cost: Assessment is free, but need to purchase materials to make kit Time: 25 minutes for CAHAI-13 version Training: None required, but a training video is available for purchase from http://www.cahai.ca/.	Cost: Free Time: 10 minutes Training: None (McDermott, 2015)	Cost: Free Time: Around 10-15 minutes Training: None
Reliability	Test-retest reliability: CAHAI-13: 0.98, CAHAI-9: 0.97, CAHAI-8: 0.97, CAHAI-7: 0.96. (Barreca, et al., 2006)	Kappa values ranging from 0.63-0.96 to 0.62-0.79. (Van de Winckel, et al., 2006) Kappa-value of >0.86 on 10 out of 17 items. (Van de Winckel, et al., 2006) SEM= 2.68. (Johansson & Häger, 2012) ICC of >0.98 for subscores and total score. (Johansson & Häger, 2012)	ICC range 0.37-.74 (Chen, et al., 2012)
Validity	Construct validity (ROC curve): CAHAI-13: 0.95, CAHAI-9: 0.94, CAHAI-8: 0.93 CAHAI-7: 0.97. (Barreca, et al., 2006)	Correlation score of $r = 0.87$ for MESUPES and the UE part of the M MAS. (Johansson & Häger, 2012)	No differential functioning found: Hierarchy of items the same across participants. Concurrent validity with FMA good (rho = 0.58-0.69), with SIS-Hand good (rho = 0.57-0.59). (Wu, et al., 2011) Predictive validity: Significant correlation with both FMA and SIS-Hand results post-treatment (rho > 0.56, $p < 0.01$) (Wu, et al., 2011)
Internal consistency	CAHAI-13: 0.98, CAHAI-9: 0.98, CAHAI-8: 0.98, CAHAI-7: 0.97. (Barreca, et al., 2006)	Not addressed	Unidimensionality found: 72.2% of variance explained by Rasch dimension (Chen, et al., 2014)
Minimal clinical significant	Not addressed.	Minimum detectable change: 18% (Johansson & Häger, 2012)	Not addressed.

difference (minimal detected change)			
Responsiveness	Not addressed	Not addressed	Responsiveness of subacute and chronic versions moderate: *d* = 0.41-0.52 (Wu, et al., 2011)
Resources	Link to administration manual: http://www.cahai.ca	Link to administration manual: http://www.strokengine.ca/wp-content/uploads/2015/09/SCALE-MESUPES-English.pdf	Link to administration of full WMFT: https://www.google.com/search?q=administering+the+wolf+motor+function+test&ie=utf-8&oe=utf-8 Link for streamlined version outline: http://www.ncbi.nlm.nih.gov/pubmed/19276293

Abbreviations Key: MESUPES- Motor Evaluation Scale for Upper Extremity in Stroke patients, CAHAI- The Chedoke Arm and Hand Activity Inventory, FMA – Fugl-Meyer Assessment, MESUPES - Motor Evaluation Scale for Upper Extremity in Stroke Patients, M MAS – Modified Motor Assessment Scale, ROC – receiver operating characteristic, SEM- standard error of the mean, SIS – Stroke Impact Scale, ICC- intraclass correlation coefficient, UE- upper extremity.

Table 8

Summary of Key Finding for Parkinson's Disease Assessments

	MDS-UPDRS	RDRS
Brief description	Developed as a comprehensive, flexible, and efficient means keep track of disability as a result of Parkinson's disease. There are 65 items that are measured using an ordinal scale of 0-4. 1) "Non-motor experiences of daily living", 2) "Motor experiences of daily living", 3) "Motor examination", and 4) "Motor complications". It combines patient report with clinician observations (Movement Disorder Society, 2008).	Patient is observed while drinking from a cup, putting on and buttoning a coat, and walking (Colosimo et al., 2010). Test has similar items as the Obeso Dyskinesia Rating Scale (Colosimo et al., 2010). Uses a rating scale (0-4) with descriptors to rate dyskinesia in terms of interference with function. In the original version, 3 activities were rated with the highest rating was entered as the disability score. Some modified versions assign 3 separate scores. Others include additional tasks such as communication (Colosimo et al., 2010).
Actual component tested/aim of measure	Motor impairment and ability to engage in ADL	Assess the severity of dyskinesia in functional contexts
Equipment Required	None	None
Clinical Utility	Cost: free (must request permission from Movement Disorder Society to use) Time: 30 minutes Training: Strongly recommended (related expenses apply)	Cost: free Time: 5 minutes or less Training: required, not specified
Reliability	Not established.	Limited reliability data; a study by Goetz et al, 1994 demonstrated significant inter-rater and intra-rater consistency (Raad, 2014).
Validity	Concurrent: Part I: Excellent (r = 0.76)	Not established.

	Part II: Excellent (r = 0.92) Part III: Excellent (r = 0.96) Part IV: Excellent (r = 0.89) (Raad, 2016)	
Internal consistency	Part I: Adequate (α = 0.79) Part II: Excellent (α = 0.90) Part III: Excellent (α = 0.93) Part IV: Adequate (α = 0.79) (Raad, 2016)	Not established.
Minimal clinical significant difference (minimal detected change)	Part I: Adequate (lowest 0.1%/highest 0.8%) Part II: Adequate (lowest 0.1%/highest 0.7%) Part III: Adequate (lowest 0.1%/highest 0.2%) Part IV: Poor (floor effect but no ceiling effect) (Raad, 2016)	Not established.
Responsiveness	Not established.	Poor; P = 0.621, effect size = 0.003 (Raad, 2014; Goetz et al., 2013)
Resources	Link to administration manual: http://www.movementdisorders.org/MDS-Files1/PDFs/MDS-UPDRS_English_FINAL.pdf	

Abbreviation Key: ADL - activities of daily living, ICC - intraclass correlation coefficient, MDS-UPDRS - Movement Disorder Society Unified Parkinson's Disease Rating Scale

ANALYSIS OF FINDINGS

Analysis of Stroke Findings

A search of two online databases resulted in five assessments that were suitable for assessing upper extremity voluntary motor control post-stroke and that fit the clinical utility requirements of the clinicians. Of the five included assessments, only the AMAT and the CAHAI are comprised of functional tasks. During our second meeting with the clinicians regarding the five assessments, they all agreed that they preferred the idea of using an assessment that involved functional tasks. The inclusion of functional tasks within the AMAT and the CAHAI makes the results more generalizable to measuring participation in meaningful activities and is more aligned with occupational therapy practice. The use of functional tasks is also shown to improve motor performance in clients post-stroke (Preissner, 2010). The structure and scoring of both the AMAT and CAHAI allow for skilled observations to occur during administration of the assessments (O'Dell et al., 2011; Barreca et al., 2004). Assessment and observation of functional tasks has the potential to provide occupational therapy clinicians with the most generalizable and useful data for creating and implementing treatment plans to address voluntary motor control and participation in valued occupations. The AMAT and CAHAI also have very strong validity, reliability, and sensitivity data to support their use with clients who have mild-to-severe physical impairments (See Table 7).

Other assessments included in this appraisal were the S-WMFT, ARAT, and MESUPES. The S-WMFT consists of a combination of functional tasks and isolated movements. The authors of this review do not recommend its use with clients who have more severe impairments following stroke due to the absence of research with that population. While the S-WMFT is very quick to administer, the limitation in populations studied make it less useful for the collaborating skilled nursing facility rehabilitation center as, generally, they provide treatment to more severely impaired patients. The psychometric data including validity, reliability, and sensitivity for the S-WMFT is just average (See Table 7). The ARAT and MESUPES both consist solely of isolated movement tasks, which is less applicable to function in ADL/IADL and does not address the concerns of the collaborating clinicians. The ARAT does have strong reliability, moderate predictive validity, strong concurrent validity, and variable-strength construct

validity (See Table 7). Additionally, the minimal clinically important difference (MCID) of the ARAT is good, falling between 10 and 30% (See Table 7). However, the strict inclusion and exclusion criteria of the ARAT-related studies, the lack of data on its application with more severe stroke patients, and the lengthy set-up time of the ARAT demonstrates low generalizability and low clinical utility, both of which outweigh its potential strengths. The MESUPES measures the quality of upper extremity movement, which is valuable to the clinicians. The psychometric data including reliability and validity is also strong; however, there is limited research and psychometric data available to make accurate comparisons with other assessments. For example, the MESUPES-related studies included in this appraisal did not investigate the sensitivity of the assessment (See Table 7).

The psychometric data found for all five assessments ranged from moderate to good. However, MCID and responsiveness data was not available for all the assessments, and future research should focus on filling the gaps in data. Taking the factors of each assessment into consideration, the authors of this CAT recommend that the collaborating clinicians utilize the CAHAI-9 with their clients who have experienced a stroke. This assessment is comprised solely of functional tasks and is appropriate for a wide range of impairment levels post-stroke. In addition, the CAHAI-9 fits the clinicians' preferences for an upper extremity assessment. It is under $300 to assemble the kit, it takes less than 30 minutes to administer and score, and it measures voluntary motor control within the contexts of functional tasks.

Analysis of PD Findings

A search of the current research on upper extremity function in patients with PD resulted in a recent systematic review that the authors of this CAT were able to use as a basis for our search. Utilizing the pragmatic parameters identified by the clinicians, two specific measures were identified that met their needs. The Rush Dyskinesia Rating Scale (RDRS) and Movement Disorder Society's Unified Parkinson's Disease Rating Scale (MDS-UPDRS) both consider the impact of dyskinesia on performing functional tasks. Occupational therapy clinicians who use either of the two assessments are able to determine how client factors or performance skills impact a client's participation and performance in occupations. As with stroke, considering only components of movement out of context does not always provide clinicians

with enough information to understand a client's participation in occupation or identify functional outcomes of the therapy. The Rush Dyskinesia Rating Scale (RDRS) involves only three functional tasks, limiting generalizability to other ADL/IADL tasks. Conversely, the MDS-UPDRS provides clinicians with a more comprehensive view of client function and limitations. Both assessments are quick to administer and are available for free, though there is an associated cost with respect to the training that is necessary for competently administering either. The clinicians would benefit most from using the MDS-UPDRS, because it provides more information regarding client performance. The MDS-UPDRS was found to have very strong validity as well (See Table 8). It is important to note that very little or no psychometric data was available for the RDRS, which made analysis between the two assessments difficult (See Table 8). It is recommended that the RDRS be researched further before being administered to clients. For clients with PD, the authors of this CAT recommend using the MDS-UPDRS. The MDS-UPDRS is more comprehensive, as it allows for a more complete assessment of the functional limitations of persons with PD.

Overall Summary

This critical appraisal of the topic provided an in-depth analysis of the research to determine measures of UE functional use that best meet the needs of occupational therapists who practice within a skilled nursing facility. The CAHAI-9 and the MDS-UPDRS are both measures that have strong psychometric properties that support their use. The psychometric data that is available for both assessments indicate that they are highly valid, reliable, and sensitive for measuring upper extremity voluntary motor control in patients with mild to severe impairments as a result of either stroke or PD. Additionally, the two identified assessments meet the specific needs of the clinicians in terms of cost, time required to administer and score, and ability to measure motor control in the context of function.

IMPLICATIONS

Consumers

The target population for this review is older adults in a skilled nursing facility who are post-stroke or with PD. This review reveals that there are assessments available that can be administered in less than 30 minutes, cost less than $300, and can be used to set therapy baselines and track progress. Ideally, it would be beneficial for a therapist to choose an assessment that is the best fit for the individual client. Consumers can use this review to see a summary of the research available on five assessments used for stroke and two for PD. They can also be more aware of what specifically is being measured by these assessments and the available data on reliability, validity, and other psychometric data. They can know that the intervention and treatment being provided by their therapist is rooted in research-driven data, and is sensitive to changes in their performance and function. The findings of this paper can also be used to justify the continuation of services by demonstrating to insurance companies and other providers that gains have been made or that the potential for functional gains is present.

Practitioners

Standardized measures are an important part of the occupational therapy process in that they provide information on functional impairments for intervention planning and help to quantify outcomes. Therapists will need to use clinical reasoning to determine which assessment is best suited to a particular population, diagnosis or stage of diagnosis (such as acute, subacute, or chronic), and level of functional impairment. For this, clinicians need reliable and valid measures. Practicing therapists can use the data in this CAT to make informed choices on which of the various assessments is a best fit for their client. Not all of the assessments included in this study have established minimal clinically important differences (MCID), which can affect a clinician's decision on whether an assessment will be useful beyond the initial evaluation. Clinicians should be certain that the assessments they are using can accurately and reliably track the changes that make a functional difference to their clients. Lower or unestablished MCID must be taken into consideration when interpreting the results of any reassessments, as it would be unclear what amount of change is important to the client or clinically significant. All of the assessments

included were screened to fit within the existing structure of the facility, and should not require any changes to the existing rehabilitation framework or program.

Researchers

Critical appraisals are useful for researchers to see the current state of research on a specific topic. In addition, appraisals can be used to guide or provide direction on future research. In general, all of the assessments mentioned in this review could benefit from further refinement. For example, some assessments would benefit from adopting interval scoring as opposed to ordinal scoring in order to improve the ease of researching psychometric data for the assessments. Some of the newer assessments need more studies that evaluate the psychometric properties and the publication of detailed standardized administration guidelines in order to ensure high reliability and promote easy access of information for clinicians. Minimum detectable change is a very important piece of psychometric data that all assessments should have, thus research should be done to ensure that data is available for each of these assessments. There are far fewer high quality studies on assessments for voluntary motor control with PD. More research on these assessments would benefit clinicians who work with that population. In particular, the Parkinson's assessments could use further review in order to determine validity and responsiveness. All of the assessments would benefit from further research conducted on populations with a wider variety of ages and levels of impairments. Qualitative studies that address the concerns of therapists and or the experience of the clients when utilizing these measures may also provide valuable information that would improve the clinical utility and acceptability of the measures.

Involvement Plan

The occupational therapists at Cottesmore of Life Care, a skilled nursing facility in Gig Harbor, WA, treat numerous clients who are status post stroke or who have been diagnosed with PD. The clinicians wanted a new measure of upper extremity functional use or abilities that could be used with these populations to provide information on initial level of function and track gains made throughout treatment. The authors of this CAT narrowed down the available outcome measures based on the clinical

utility requirements of the collaborating clinicians and provided in depth information on five assessments for stroke and two for PD.

Developing an Involvement Plan

To develop an involvement plan that would meet the needs of the collaborating clinicians, the authors of this CAT met with them to discuss our results and to elicit their input on what knowledge translation activities would meet their needs. When the authors of this CAT initially presented our findings to the collaborating clinicians, they liked the functional tasks that formed part of the Motor Evaluation Scale for Upper Extremity in Stroke Patients (MESUPES), as they felt that the tasks represented an accurate assessment of motor control return. This initial interest may have come from the fact that, no list of tasks for the other assessments was provided to allow for comparison. In future meetings, the authors of this CAT made sure to include a list of task items for all assessments. The collaborating clinicians felt that the tasks represented an accurate assessment of motor control return. They were concerned that the Chedoke Arm and Hand Activity Index (CAHAI) would not accurately assess function in clients with more severe stroke, and would not be sensitive enough to track changes. In subsequent meetings, the authors of this CAT addressed that concern by explaining that a large portion of the research on the CAHAI has demonstrated it is appropriate for use with patients who have moderate to severe impairments post-stroke. We described the tasks and task objects that the CAHAI encompasses, explaining that many of the tasks within the assessment can be performed bilaterally. Even individuals with severe hemiparesis, for example, should be able to experience some degree of success. After this concern was addressed, the clinicians agreed that the CAHAI was a good fit for their setting and populations.

The authors of this CAT informed the clinicians that finding research on assessments that measure upper extremity functional use for patients with Parkinson's disease had been more challenging. We recommended the Movement Disorder Society's Unified Parkinson's Disease Rating Scale (MDS-UPDRS), mainly due to the comprehensive information that could be derived from its administration. One of the clinicians at the SNF reported that patients who have Parkinson's disease are not referred to

occupational therapy at their clinic unless it is a secondary condition that impacts a primary condition. There was some miscommunication early on in the process between the collaborating clinicians and authors of this CAT regarding the facility's client population and needs. It was unknown whether they will consider implementing an assessment specifically designed for use with Parkinson's disease as we recommended. We noted in our brochure that the MDS-UPDRS is a useful assessment for all aspects of PD, that it is used to help assess ADL, mentation, and PD symptoms, and can be used to track clients' progress over time.

The authors of this CAT informed the clinicians that the next phase of the project is the implementation phase and that there are several ways that we could proceed. We created time for the clinicians to see the assessments and hear explicitly how they are scored, which may increase the likelihood that the assessment will be adopted. Reading a manual or brochure does not convey the complexity of administering these measures or opportunities for skilled observation provided by the demonstration of the CAHAI. An interactive learning session may also reduce some of the barriers to future use by the clinicians. While it was not feasible to create a test kit for the facility, we provided very detailed instructions on how a kit is created (see Appendix D), a cost estimate for creating the kit, and then showed the clinicians a completed kit. Providing this information may increase the chance of the facility creating and using a kit for the assessment. The clinicians were most interested in having us bring the CAHAI and demonstrate its administration on either one of them or on one of their clients. We provided the clinicians with a cost estimate for assembling the kit. To help the clinicians determine if they would like to proceed with adoption of an assessment for patients with Parkinson's disease, we also agreed to continue to research the MDS-UPDRS. We provided the clinicians with a brochure that outlined where they could obtain the manual and the necessary training. It was anticipated that this brochure would also help the clinicians to understand how the MDS-UPDRS is applicable to their population and provides valuable information on various aspects of PD. At the end of our meeting, the clinicians planned to read and share the custom table provided to them so that they could learn more about each of the assessments

included in our CAT. Providing our table in a more reader-friendly format such as a booklet also formed part of the involvement plan.

Potential Barriers to Knowledge Translation

The clinicians were enthusiastic about adopting use of the CAHAI. Possible barriers to implementation that we anticipated include a lack of time dedicated to learning the new assessment or a lack of administrative support for change. Given the high productivity requirements of this facility, it could be very difficult to get the OTR and OTA staff together to provide education on a new or unfamiliar assessment, or to make time to problem solve any future concerns or issues that arise in regards to the assessment. Other members of the rehabilitation team may not see the value of the new assessment or how it will enhance overall service provision. On a personal level, a clinician who does not agree with our recommendation may be much more reluctant to learn and administer the assessment with their clients, or may see it as less valuable than other assessments and choose to allocate time in other ways than becoming familiar with the measure. Such attitudinal and systemic barriers can be addressed by identifying the root of the resistance to change with surveys or more informal methods such as conversations (McCluskey, 2013). The authors of this CAT used the in-service as a way to address potential barriers in the rehabilitation staff by providing information on the administration and adoption of the measures, and answering questions about its use. The involvement of the rehabilitation director in the process also reduced potential systemic barriers such as allocation of time for implementation.

Considering the enthusiasm demonstrated throughout the knowledge translation process to adopt use of the CAHAI and the fact that our group carefully selected assessments based on clinical utility factors addressed by the clinicians such as time and cost to administer, the only barriers to implementing a new assessment should be the time required to learn and practice administering the assessment, creating and maintaining the test kit, and providing educational information to future clinicians. However, there are many more barriers that could prevent the clinicians from adopting the use of the MDS-UPDRS. The clinicians do not currently recognize the value of using the MDS-UPDRS, likely as a result of the limited information that we provided them. Limited knowledge of the MDS-UPDRS, potential mismatch with the

clinicians' needs (such as not treating PD as a primary diagnosis or not fitting with their perceived evaluation needs), and time are all potential barriers to adopting use of the assessment.

The Involvement Plan

Knowledge translation involves the process of taking evidence and implementing it in clinical care with clients. Occupational therapists and other health care professionals are expected to provide care guided by best practices to which knowledge translation is closely tied. This critical appraisal of a topic (CAT) provided a more in-depth analysis of the psychometric properties of each assessment as they were administered to specific populations. It is important to consider how different assessments may be more or less appropriate for particular client populations. For example, the sensitivity of some assessments may differ depending on the severity of a person's health complications. By working directly with practicing clinicians and addressing the specific needs of their facility, the authors of this CAT were able to bridge the gap between the published research and implementation of a new assessment.

The first step of translating knowledge into practice was to present our findings to the clinicians in a way that is easily readable and understandable, and that clearly links our findings to their needs. We presented the clinicians with an electronic copy of our full paper, printed summaries of our findings (Tables 7 & 8), and the implications of our findings on various audiences. The abbreviated printed version reduced the amount of reading necessary to understand the main points of this appraisal, while the full electronic version provided them with access to the in-depth research behind our recommendations.

The second stage of translation involved the creation of two fliers to present to the clinicians. One of these fliers highlighted the strengths of the Chedoke Arm and Hand Activity Index, 9-item version (CAHAI-9), which was our recommendation for use with clients who are post-stroke. The flier (see Appendix B) is a brief summary of the information provided in our paper and in the administration manual for the assessment. We also printed and bound a copy of the administration manual to present to the facility with a list of estimated costs to create their own test kit. The second flier (see Appendix C) provided information on our recommendation for clients with Parkinson's disease (PD). This flier focused more on general information about the assessment, and included much more information on how to

acquire the assessment. Because the clinicians seemed less inclined to adopt an assessment designed specifically for PD at their facility, the flier also contained information on the merits of treating PD as a primary diagnosis and on collecting data on occupational therapy for upper extremity motor control. Such data collection could help to reduce the shortage of information available in this field.

In order to help the clinicians better understand and feel comfortable and confident with the CAHAI-9, we presented the assessment as an in-service. One of the authors of this CAT administered the assessment on a second author. The person simulating the client demonstrated various levels of impairment and the clinicians discussed and scored the performance according to the guidelines in the administration manual. By having the clinicians score a 'client' while we were there, we were able to determine if they are having difficulty scoring the measure and clarified administration procedures as needed. At this time, the authors of this CAT presented them with both fliers, the test manual, and the list of items and costs. We brought a completed test kit from the University of Puget Sound as an example, and used this to demonstrate the administration of the assessment. Three test items were demonstrated, and were then scored collaboratively. This collaborative scoring process gave the clinicians a chance to be involved in the process, to talk over the scoring guidelines and ask questions about specific movements or cuing. Following the in-service, the authors of this CAT led the clinicians in a discussion of the relative strengths and weaknesses of the assessment, and talked over potential opportunities and barriers to implementing it in their facility. Responses to the strengths, weaknesses, opportunities, and threats (SWOT) analysis were recorded by the students. The steps taken to complete this involvement plan are outlined in Table 9.

Table 9

Tasks, Dates, and Steps Involved in Implementation Plan

Task/Product	**Deadline Date**	**Completion Date**	**Steps to achieve the final outcome**
Send an email to the rehabilitation director explicitly stating that our paper will be undergoing revisions	3/16/2016	3/16/2016	1. Review task objects for the Chedoke and check prices in the community.

over the duration of the implementation phase. This email will also outline our implementation plan and confirm that they are interested in having us put a kit together with a total cost estimate.			2. Decide who will write the email. 3. Write and send the email.
Provide copy of our summary table and implications	3/26/2016	3/26/2016	1. Provide summary of findings 2. Provide summary of implications of results
Create a flier about how to find manual and set-up a test kit for CAHAI-9	4/1/2016	4/1/2016	1. Find links to manual and test items needed for the kit. 2. Create and print the flier.
Create a brochure with information on training resources and more in-depth information about the MDS-UPDRS	4/5/2016	4/5/2016	1. Continue to research UPDRS to be sure the assessment meets the clinicians' needs and find links to training resources and manual. 2. Design and print the brochure.
Create administration manual	4/11/2016	4/11/2016	1. Create binder with test creation info, administration manual, and instructions for making test items per guidelines 2. Use manual in demonstration and leave with the clinicians as additional resource
Take CAHAI-9 test kit and demonstrate administration of test	4/14/2016	4/21/2016	1. Practice administering the Chedoke. 2. Arrange a time to present to clinicians. 3. Demonstrate the Chedoke on one of the researchers 4. Grade the Chedoke collaboratively with the clinicians 5. Respond to questions that arose in the demonstration
Analyze CAHAI-9 with clinicians	4/14/2016	4/21/2016	1. Use SWOT form to identify strengths, weaknesses, opportunities, and barriers
Follow-up with clinicians	4/19/2016	4/26/2016	1. Send survey to clinicians to ask if any thoughts/concerns came up after the presentation 2. Ask if there is any further help we can provide given our

			timeline 3. Ask them to think about possible research topics for next year

Monitoring Outcomes of Knowledge Translation

The outcomes of the above activities were monitored by having the clinicians provide feedback on how helpful each activity was for them. Throughout the in-service, we asked for feedback on the assessment and the administration process. We used SWOT analysis with the clinicians after completion of the in-service to further identify specific concerns or strengths they identified during the demonstration and grading. Some questions were pre-identified, but the discussion flowed from their thoughts and comments. At the time of the in-service, we reminded them of the opportunity for future collaboration with students from the University of Puget Sound on similar research topics. Additionally, a survey was sent to the facility after the in-service asking how useful the in-service was, if any additional questions have come up, and to check if any further assistance could be provided at this time. A copy of the survey is included in Appendix E.

Products Created

The clinicians were provided with a brochure covering a general overview of the CAHAI (Appendix B). A more comprehensive brochure on the MDS-UPDRS was also provided (Appendix C), which gave an outline of the measure, its psychometric data, and arguments in favor of its use. Other documents provided to the clinicians included a copy of our summary tables (Tables 7 & 8), implications from the CAT, and a bound copy of the CAHAI administration manual. A list of items (see Appendix A) needed for the assessment and projected costs was also included in the copy of the administration manual.

Effectiveness of the Tasks and Products

Effectiveness of the In-Service

On April 21st, 2016, the authors of this CAT visited the clinicians at Cottesmore for a third time. We feel that this meeting went well and that the clinicians were very receptive to our ideas. During this meeting, we performed a demonstration of three task items of the Chedoke Arm and Hand Activity Inventory (CAHAI). We then scored the assessment with the clinicians and talked over questions regarding scoring. One clinician had some concerns about the scoring and we were able to answer her questions and help her feel more comfortable about the scoring. This interactive way of teaching the clinicians about the CAHAI was effective because it gave them an opportunity to ask us questions in the moment regarding administration and scoring. By the conclusion of the activity, the clinicians seemed to have a clear understanding of how the scoring and administration works and many of their concerns were addressed.

After concluding the demonstration, we went over the assessment with the clinicians and performed a SWOT analysis (see Appendix A). The SWOT analysis revealed that the clinicians like the functional, relevant nature of the tasks, and that the familiar nature of the tasks will help to address concerns about cognition. Most clients will have a habitual response to tasks like dialing a phone or cutting food, and will be more able to complete a familiar task than a novel one. The clinicians also identified the robust research and wide variety of stroke severity in the populations studied as a strength. The clinicians were concerned that use of the assessment might not be compatible with quantifying patients' severity of impairment for Medicare G-codes, which they use frequently. They also had questions about the bilateral nature of the tasks, when some clients may be more familiar performing them one-handed. We addressed this by reiterating that the test asks clients to perform tasks bilaterally, and included cuing for this in the administration instructions. The clinicians were also concerned that there would be a big learning curve for this assessment, and that it might be difficult to get all of the therapists and therapy assistants feeling competent on administration. The CAHAI has excellent inter-rater reliability, but some sort of standardization will be necessary within the facility. The clinicians felt

that the assessment would track changes in patients' functional level adequately and that its sensitivity would increase as a clinician developed a relationship and built trust with his or her client. The clinicians mentioned that even if this assessment was not included on the initial evaluation, it would provide valuable information for treatment.

Effectiveness of the CAHAI Brochure

A brochure giving a brief overview of the assessment was provided to the clinicians, which reiterated some of the key strengths of the assessment that led to our recommendation for its use. The clinicians were receptive to the information and felt that it would be beneficial to provide this to therapists who were unable to attend the in-service and to clinicians who are hired at the facility in the future. The brochure would be given alongside more in-depth training on the use of the CAHAI.

Effectiveness of the MDS-UPDRS Brochure

The authors of this CAT also presented the clinicians with a brochure about the MDS-UPDRS and its merits. The brochure gave a brief overview of the assessment and a summary of the psychometric data. It also covered the value of using diagnosis-specific assessments and recommended the adoption of the MDS-UPDRS for use with their clients with PD. The clinicians were initially hesitant about adoption of an assessment for PD, but responded positively to the brochure and seemed to better see the value of implementation.

Analysis of In-service

Overall, the meeting went well and ended on a positive note. The clinicians all agreed that they would have to take some time to work out the details of administering the CAHAI-9 with their patients, but as a whole they were in agreement that they would like to start utilizing it in their clinic. The rehabilitation director told the other clinicians that she would like to set up a meeting with them soon to discuss it further and take the next steps necessary to use the assessment in their clinic.

Results of Follow-Up Survey

When asked how likely they were to implement the assessments in their facility, the clinicians rated the CAHAI an 8/10 and the MDS-UPDRS a 7/10, indicating that it is likely that they will implement

them. They indicated that the assessment was very informative and that the research regarding the assessments was ‘extremely useful’ to their setting. When asked, the clinicians did not have any suggestions for how the research and process of knowledge translation could be improved upon and they indicated that the process utilized was efficient and helpful.

The authors of this CAT feel that the products and tasks we completed for our research project and presented to the clinicians were effective in portraying our results in a concise and timely manner.

Analysis of the Overall Process

In September, 2015, we were assigned to work with Nan, who is an occupational therapist and the rehabilitation manager at Cottesmore of Life Care skilled nursing facility in Gig Harbor, WA. Through meeting with the occupational therapists at Cottesmore, we collectively chose a topic to research and commenced our evidence-based research journey.

Beginning the process was difficult, as we were unsure how best to approach the topic without being overwhelmed with information. We had no idea that the scope of the assigned topic would be so large, and found it very hard to develop a starting point for all of the information available. Our project chair provided valuable insight in how to first narrow the number of assessments, then look only at those specific assessments.

When we initially began searching for articles, it was difficult to keep track of all the articles and categorize them into why we did not want to utilize certain articles in our CAT table. After creating our CAT table, we analyzed our findings and made conclusions. It was difficult to compare results given that each article covered different pieces of psychometric data. However, we tried as best we could to piece the information together to understand which assessments had higher psychometric data and why. As we began to write the analysis of findings, implications and conclusion, we found it difficult to not become redundant with the information we presented.

Once we had a rough draft of our paper, we met again with the clinicians at the SNF. This meeting was helpful because we were able to present our findings and ensure that they had an interest in some of the assessments we were analyzing. It was a good chance for them to ask us questions and for us

to further understand what their priorities were for an assessment. During this meeting we did feel a bit unprepared, as we had spent so much time finding data that our analyses were very primitive.

The biggest challenge overall of this process was that we were required to search for assessments for two separate populations. This made the scope of our paper rather large and entailed developing two separate search strategies due to the lack of overlap between the two populations. Addressing two populations created challenges related to formatting our paper in an organized manner as well.

Though we encountered many challenges with the process of researching and writing the paper, our final meeting with the clinicians went very well. They were impressed with our knowledge and amount of research. They were sincerely interested in using the assessments we recommended in their clinic. The rehabilitation director said that she would meet with her team later on to further discuss the implementation of the assessment in their clinic. Hearing the clinician's positive feedback and excitement with our findings was very encouraging and made us feel empowered as future clinicians and as research analysts.

Recommendations for Follow-Up Projects

During our initial meeting, the clinicians suggested numerous topics and questions for further research. One of the ideas initially suggested involved research on effective home exercise programs. They would like to come up with the most appropriate home exercise program to send home with their patients.

Future research within the topic of upper extremity (UE) assessments could include researching all available UE assessments for clients with stroke and comparing the assessments against each other. Each assessment could be ranked by cost, time to administer and psychometric data. This sort of research would provide clinicians with a good overview of all available assessments and allow them to choose an assessment to use based on their specific set of criteria.

Further research can also be done to determine potential upper extremity PD-specific assessments used within occupational therapy. There are currently very few PD assessments that are appropriate for

use in OT. Of those that are available, little research has been reported on their psychometric properties and the value that these assessments provide to the OT process.

CONCLUSION

The occupational therapy clinicians at the skilled nursing facility with whom the authors of this CAT collaborated currently use the Modified Barthel Index for clients with stroke and PD, but find the measure to be inadequately sensitive to changes in voluntary motor control in the upper extremities. The authors of this critical appraisal of the topic assisted the clinicians in identifying appropriate reliable, valid, and sensitive assessments based on analysis of the evidence with consideration of the practicality of administering each of the measurement tools within the preexisting framework of the facility. A critical appraisal of the topic was necessary to ethically and efficiently initiate the knowledge translation process on behalf of these working clinicians.

There are many upper extremity voluntary motor control assessments available for use by occupational therapists. Of the assessments available, five assessments for stroke and two assessments for PD were found that fit the clinical utility needs of the clinicians. All of the stroke assessments had robust psychometric data of a high enough quality to be beneficial to the clinicians at a skilled nursing facility rehabilitation clinic. Less data was available for the PD assessments: only four assessments had reliability data and only one had validity data.

Based on these findings, the authors recommend the use of the CAHAI-9 with patients who are post-stroke. For PD, we recommend the MDS-UPDRS. However, the authors of this CAT emphasize that therapists must use clinical reasoning to determine if a given assessment will be the best fit for their clients. With clients post stroke, the severity and presence of comorbidities may affect the results of the assessment, and psychometric data may not be as strong for some populations as for others. For PD, the lack of information available and the need for clinician training are the biggest drawbacks to the assessments included. Further research on upper extremity function in PD would benefit the profession and increase the strength of all available assessments.

Knowledge translation is a complex process that requires more than reading peer-reviewed journal articles (Metzler & Metz, 2010). Possible barriers to the implementation of research knowledge into practice include the large volume of research literature available and the often difficult process of critically evaluating the evidence (Metzler & Metz, 2010). Rehabilitation professionals face many obstacles to knowledge translation that are unique to their niche in healthcare. For instance, rehabilitation practice would benefit from evidence derived from case studies to randomized controlled trials and every type in between, meaning a therapist could more readily find evidence from a study that lacks rigor. Additionally, there may be more opportunities for a therapist to misinterpret or incorrectly apply an intervention because there are often specific conditions, such as the intensity and frequency in which an intervention was applied, of the protocol used in studies (Metzler & Metz, 2010). Likewise, the sample characteristics such as socioeconomic status, culture, and impairment level may differ from the population treated by a therapist. Finally, rehabilitation clinicians do not always have the same resources available to them that are used by researchers (Metzler & Metz, 2010). One of the aims by the authors of this critical appraisal of the topic was to streamline this complex process for the collaborating clinicians so that the likelihood of adopting and using more appropriate and sensitive assessments would increase. The findings of this study may be extrapolated to other clinicians who work in skilled nursing facilities that serve a similar client population.

References

Asher, E. (2014). *Asher's occupational therapy assessment tools: An annotated index*. White Plains, MD: AOTA Press.

Barreca, S., & Masters, L. (2004). *The Chedoke Arm and Hand Activity Inventory* [Measurement instrument]. Retrieved from www.cahai.ca

Barreca, S., Stratford, P., Lambert, C., Masters, L., & Streiner, D. (2005). Test-retest reliability, validity, and sensitivity of the Chedoke Arm and Hand Activity Inventory: A new measure of upper-limb function for survivors of stroke. *Archives of Physical Medicine and Rehabilitation, 86,* 1616-1622.

Barreca, S. R., Stratford, P. W., Masters, L. M., Lambert C. L., & Griffiths, J. (2006). Comparing two version of the Chedoke Arm and Hand Activity Inventory with the Action Research Arm Test. *Physical Therapy, 86,* 245-253.

Barreca, S., Stratford, P., Masters, L., Lambert, C., Griffiths, J., & McBay, C. (2006). Validation of three shortened versions of the Chedoke Arm and Hand Activity Inventory. *Physiotherapy Canada, 58(2),* 148-156.

Bogard, K., Wolf, S., Zhang, Q., Thompson, P., Morris, D., & Nichols-Larsen, D. (2009). Can the Wolf Motor Function Test be streamlined? *Neurorehabilitation and Neural Repair, 23,* 422-428. http://dx.doi.org/10.1177/1545968308331141

Center on Knowledge Translation for Disability and Rehabilitation Research [KTDRR]. (2005). Retrieved from http://ktdrr.org/ktlibrary/articles_pubs/ncddrwork/focus/focus10/Focus10.pdf

Chen, H., Lin, K., Wu, C., & Chen, C. (2012). Rasch validation and predictive validity of the Action Research Arm Test in patients receiving stroke rehabilitation. *Archives of Physical Medicine and Rehabilitation, 93,* 1039-1045. http://dx.doi.org/10.1016/j.apmr.2011.11.033

Chen, H., Wu, C., Lin, K., Chen, H., Chen, C., & Chen, C. (2012). Rasch validation of the Streamlined Wolf Motor Function Test in people with chronic stroke and subacute stroke. *Physical Therapy, 8,* 1017-1026.

Chen, H., Wu, C., Lin, K., Jang, Y., Lin, S., Cheng, J., Chung, C., & Yan, Y. (2014). Measurement properties of the Streamlined Wolf Motor Function Test in patients at subacute to chronic stages after stroke. *Neurorehabilitation and Neural Repair, 28,* 839-846. http://dx.doi.org/1545968314526643

Colosimo, C., Martinez-Martin, P., Fabbrini, G., Hauser, R., Merello, M., Miyasaki, J., … Goetz, C. (2010). Task force report on scales to assess dyskinesia in Parkinson's disease: Critique and recommendations. *Movement Disorders, 25,* 1131-1142.

Figueiredo, S. (2009). Chedoke Arm and Hand Activity Inventory (CAHAI). In Nicole Korner-Bitensky & Elissa Sitcoff (Eds). *Stroke Engine: Assessments- physical upper extremity.* Montreal: Canadian Project for Stroke Recovery. Retrieved February 25, 2016, from http://www.strokengine.ca/assess/cahai/

Goetz, C., Stebbins, G., Chung, K., Hauser, R., Miyasaki, J., Nicholas, A., … Song, G. (2013). Which dyskinesia scale best detects treatment response? *Movement Disorders, 28,* 341-346.

Gustafsson, L., Turpin, M., & Dorman, C. (2010). Clinical utility of the Chedoke Arm and Hand Activity Inventory for stroke rehabilitation. *Canadian Journal of Occupational Therapy, 77,* 167-173. http://dx.doi.org/10.2182/cjot.2010.77.3.6

Heinemann, A. (2010). *Rehab Measures Database.* Retrieved from http://www.rehabmeasures.org/rehabweb/allmeasures.aspx?PageView=Shared.

Heldman, D., Giuffrida, J., Chen, R., Payne, M., Mazzella, F., Duker, A. … & Espay, A. (2011). The modified bradykinesia rating scale for Parkinson's disease: Reliability and comparison with kinematic measures. *Movement Disorders, 26,* 1859-1863.

Hsieh, C.-L., Hsueh, I.-P., Chiang, F.-M., & Lin, P.-H. (1998). Inter-rater reliability and validity of the Action Research Arm Test in stroke patients. *Age and Aging, 27,* 107-113.

Hsieh, Y., Wu, C., Lin, K., Chang, Y., Chen, C., & Liu, J. (2009). Responsiveness and validity of three outcome measures of motor function after stroke rehabilitation. *Stroke, 40,* 1386-1391. http:dx.doi.org/10.1161/STROKEAHA.108.530584

Johansson, G. M., & Häger, C. K. (2012). Measurement properties of the Motor Evaluation Scale for Upper Extremity in stroke patients (MESUPES.) *Disability & Rehabilitation, 34,* 288-294. http://dx.doi.org/10.3109/09638288.2011.606343

Koh, C.-L., Hsueh, I.-P., Wang, W.-C., Sheu, C.-F., Yu, T.-Y., Wang, C.-H., & Hsieh, C.-L. (2006). Validation of the Action Research Arm Test using item response theory in patients after stroke. *Journal of Rehabilitation Medicine, 38,* 375-380. http://dx.doi.org/10.1080/16501970600803252

Kopp, B., Herta Flor, A., Platz, T., Karl-Heinz Mauritz, U., Gresser, K., McCulloch, K., & Taub, E. (1997). The arm motor ability test: Reliability, validity, and sensitivity to change of an instrument for assessing disabilities in activities of daily living. *Archives of Physical Medicine and Rehabilitation*, 78, 615-620.

Korner-Bitensky, N., & Rochette, A. (n.d.). About stroke engine. Retrieved from http://www.strokengine.ca/about-stroke-engine/

Lang, C., Bland, M., Bailey, R., Schaefer, S., & Birkenmeier, R. (2013). Assessment of upper extremity impairment, function, and activity after stroke: Foundations for clinical decision making. *Journal of Hand Therapy, 26,* 104-115.

Lang, C. E., Edwards, D. F., Birkenmeier, R. L., & Dromerick, A. W. (2008). Estimating minimal clinically important differences of upper-extremity measures early after stroke. *Archives of Physical Medicine and Rehabilitation, 89,* 1693-1700. http://dx.doi.org/10.1016/j.apmr.2008.02.022

Lang, C. E., Wagner, J. M., Dromerick, A. W., & Edwards, D. F. (2006). Measurement of upper-extremity function early after stroke: Properties of the Action Research Arm Test. *Archives of Physcial Medicine and Rehabilitation, 87,* 1605-1610. http://dx.doi.org/10.1016/j.apmr.2006.09.003

Lannin, N. (2004). Reliability, validity and factor structure of the upper limb subscale of the Motor Assessment Scale (UL-MAS) in adults following stroke. *Disability and Rehabilitation, 26(2),* 109-115.

Li, K.-Y., Lin, K.-C., Wang, T.-N., Wu, C.-Y., Huang, Y.-H., & Ouyang, P. (2012). Ability of three motor measures to predict functional outcomes reported by stroke patients after rehabilitation. *NeuroRehabilitation, 30,* 267-275. http://dx.doi.org/10.3233/NRE-2012-0755

Lin, J.-H., Hsu. M.-J., Sheu, C.-F., Wu, T.-S., Lin, R.-T., Chen, C.-H., & Hsieh, C.-L. (2009). Psychometric comparisons of 4 measures for assessing upper-extremity function in people with stroke. *Physical Therapy, 89,* 840-850. http://dx.doi.org/10.2522/ptj.20080285

McCluskey, A. (2013). Implementing evidence into practice. In Hoffman, T., Bennett, S., & Del Mar, C. (Eds.), *Evidence-based practice across the healthcare professions* (2nd ed.; 370-390). Chatswood, Australia: Elsevier

McDermott, A. (2015). Motor Evaluation Scale for Upper Extremity in Stroke Patients (MESUPES). In Annie Rochette (Ed) Stroke engine: Assessments- physical upper extremity. Montreal: Canadian Stroke Recovery Project. Retrieved February 25, 2016, from http://www.strokengine.ca/assess/mesupes/

Metzler, M. J., & Metz, G. A. (2010). Translating knowledge to practice: An occupational therapy perspective. *Australian Occupational Therapy Journal, 57,* 373-379. http://dx.doi.org/10.1111/j.1440-1630.2010.00873.x

Miller, K., Slade, A., Pallant, J., & Galea, M. (2010). Evaluation of the psychometric properties of the upper limb subscales of the Motor Assessment Scale using a Rasch analysis model. *Journal of Rehabilitation Medicine, 42,* 315-322.

Movement Disorder Society. (2008). *Movement Disorder Society's Unified Parkinson's Disease Rating Scale.* Milwaukee, WI: Movement Disorder Society.

Movement Disorder Society Task Force on Rating Scales for Parkinson's Disease. (2003).The unified Parkinson's disease rating scale (UPDRS): Status and recommendations. *Movement Disorders, 18,* 738-750.

Murphey, M., Resteghini, C., Feys, P., & Lamers, I. (2015). An overview of systematic reviews on upper extremity outcome measures after stroke. *BMC Neurology*, *15*(29), 1-15.

Ng, A. K. Y., Leung, D. P. K., & Fong, K. N. K. (2008). Clinical utility of the Action Research Arm Test, the Wolf Motor Function Test and the Motor Activity Log for hemiparetic upper extremity functions after stroke: A pilot study. *Hong Kong Journal of Occupational Therapy, 18,* 20-27. http://dx.doi.org/10.1016/s1569-1861(08)70009-3

Nijland, R., van Wegen, E., Verbunt, J., van Wijk, R., van Kordelaar, J., & Kwakkel, G. (2010). A comparison of two validated tests for upper limb function after stroke: The Wolf Motor Function Test and the Action Research Arm Test. *Journal of Rehabilitation Medicine, 42,* 694-696. http://dx.doi.org/10.2340/16501977-0560

Nordin, Å., Murphy, M. A., & Danielsson, A. (2014). Intra-rater and inter-rater reliability at the item level of the Action Research Arm Test for patients with stroke. *Journal of Rehabilitation Medicine, 46,* 738-745. http://dx.doi.org/10.2340/16501977-1831

O'Dell, M. W., Kim, G., Rivera, L., Fieo, R., Christos, P., Polistena, C., Fitzgerald, K., & Gorga, D. (2013). A psychometric evaluation of the arm motor ability test. *Journal of Rehabilitation Medicine, 45,* 519-527.

O'Dell, M., Kim, G., Finnen, L., & Polistena, C. (2011). Clinical implications of using the Arm Motor Ability Test in stroke rehabilitation. Archives *of Physical Medicine and Rehabilitation, 92,* 830-836. http://dx.doi.org/10.5014/ajot.2010.08026

Poole, J., & Whitney, S. (2001). Assessments of motor function post stroke: A review. *Physical & Occupational Therapy in Geriatrics, 19,* 1-22.

Preissner, K. (2010). Use of the Occupational Therapy Task-Oriented Approach to optimize the motor performance of a client with cognitive limitation. *The American Journal of Occupational Therapy, 64*, 727-734.

Proud, E., Miller, K., Bilney, B., Balachandran, S., McGinley, J., & Morris, M. (2015). Evaluation of measures of upper limb functioning and disability in people with Parkinson Disease: A systematic review. *Archives of Physical Medicine and Rehabilitation, 96,* 540-551.

Raad, J. (2016). Rehabilitation measures database: Arm Motor Ability Test. Retrieved from: http://www.rehabmeasures.org/Lists/RehabMeasures/DispForm.aspx?ID=1188&Source=http%3A%2F%2Fwww%2Erehabmeasures%2Eorg%2Frehabweb%2Fallmeasures%2Easpx%3FPageView%3DShared

Rehabilitation Measures Database. (n.d.). Welcome to the rehabilitation measures database. Retrieved from http://www.rehabmeasures.org/rehabweb/rhaboutus.aspx

Rowland, T., & Gustafsson, L. (2008). Assessments of upper limb ability following stroke: A review. *British Journal of Occupational Therapy, 71*, 427-437.

Rowland, T., Gustafsson, L., Turpin, M., Henderson, R., & Read, S. (2011). Chedoke Arm and Hand Activity Inventory-9 (CAHAI-9): A multi-centre investigation of clinical utility. *International Journal of Therapy and Rehabilitation, 18*(5), 290-298.

Van Delden, A. E. Q., Peper, C. E., Beek, P. J., & Kwakkel, G. (2013). Match and mismatch between objective and subjective improvements in upper limb function after stroke. *Disability and Rehabilitation, 35,* 1961-1967. http.//dx.doi.org/10.3109/09638288.2013.768303

Van de Winckel, A., Feys, H., van der Knapp, S., Messerli, R., Baronti, F., Lehmann, R., … & De Weerdt, W. (2006). Can quality of movement be measured? Rasch analysis and inter-rater reliability of the Motor Evaluation Scale for Upper Extremity in Stroke Patients (MESUPES.) *Clinical Rehabilitation, 20,* 871-884. http://dx.doi.org/ 10.1177/0269215506072181

Van der Lee, J. H., De Groot, V., Beckerman, H., Wagenaar, R. C., Lankhorst, G. J., & Bouter, L. M. (2001). The intra- and inter-rater reliability of the Action Research Arm Test: A practical test of

upper extremity function in patients with stroke. *Archives of Physical Medicine and Rehabilitation, 82,* 14-19. http://dx.doi.org/10.1053/apmr.2001.18668

Van der Lee, J. H., Roorda, L. D., Beckerman, H., Lankhorst, G. J., & Bouter, L. M. (2002). Improving the Action Research Arm Test: A unidimensional hierarchical scale. *Clinical Rehabilitation, 16,* 646-653. http://dx.doi.org/10.1191/0269215502cr534oa

Wu, C., Fu, T., Lin, K., Feng, C., Hsieh, K., Yu, H., Lin, C., Hsieh, C., & Ota, H. (2011). Assessing the Streamlined Wolf Motor Function Test as an outcome measure for stroke rehabilitation. *Neurorehabilitation and Neural Repair, 25,* 194-199. http://dx.doi.org/10/1177/1545968310381249

Appendix A

Analysis of strengths, weaknesses, opportunities and threats to implementing the Chedoke Arm and Hand Activity Inventory in the skilled nursing facility

STRENGTHS • What do you like about the CAHAI-9 • What kinds of information are you getting?	**WEAKNESSES** • What do you dislike about the measure? • Do you see areas for improvement?
-tasks related to ADL -more sensitive than just min/mod A -bilateral test, also observation -covers range of severity -basic, fine tasks • not new, will help w/ cognition	-instructions more specific • confusing if people have their typ. way is diff than what test scores -explain to clients that you are specifically looking at bilat. func so have them do it as you demo. -not in g-codes?
OPPORTUNITIES • What would this measure bring to your therapy? • What information are you getting?	**THREATS** • Would this measure detract from you services? • Would it negatively impact your facility?
-maybe not on initial eval -more sensitive once you know/ build up trust w/ client -good for Med A acute pts -excellent to track changes	-big learning curve • to feel competent -inter-rater rel. • incorp OTR/ OTA

Appendix B

Brochure on the Chedoke Arm and Hand Activity Inventory (CAHAI)

Chedoke Arm and Hand Activity Inventory (CAHAI)

-An assessment of bilateral UE use in post-STROKE patients-

The CAHAI assessment has four versions: the 7, 8, 9 and 13 task item versions. The 9-item version is recommended due to high psychometric data and length of administration time. The assessment consists of various occupation based tasks such as using the phone, opening a jar, and buttoning a shirt. The assessment is used to measure bilateral UE use and is also a good indicator of cognitive functioning.

Facts:

- Administration time: ~30 minutes
- Assessment is free- just need to assemble a kit
- Population: Any patient post- STROKE
- Developed by OT's (contact info available)
- High psychometric data
- Available online at www.cahai.ca

Tasks for (9-item version)

- Open jar of coffee
- Call 911
- Draw line with a ruler
- Pour a glass of water
- Wring out a washcloth
- Do up five buttons
- Dry back with towel
- Put toothpaste on toothbrush
- Cut putty

Appendix C

Brochure on the Movement Disorder Society's Unified Parkinson's Disease Rating Scale

Movement Disorder Society's Unified Parkinson's Disease Rating Scale (MDS-UPDRS)

-An assessment of UE use in patients with Parkinson's Disease-

The MDS-UPDRS was published in 2008. It was designed to address weaknesses, ambiguities, and areas in need of inclusion, based on research evidence, of the original UPDRS. There are four parts that include: 1) "Non-motor experiences of daily living", 2) "Motor experiences of daily living", 3) "Motor examination", and 4) "Motor complications". There are 65 items and a scale between 0-4 is used to score each item. The assessment takes approximately 30 minutes to administer.

Comparing the MDS-UPDRS to the UPDRS

The MDS is an updated and psychometrically superior version of the UPDRS that tends to focus on the impact rather than the presence of symptoms. There are 9 new items in the MDS-UPDRS including fatigue, anxious mood, urinary problems, constipation, dopamine dysregulation syndrome, participating in hobbies, toe tapping, getting in and out of bed, and freezing. The MDS-UPDRS measures lightheadedness using a 0-4 rating scale as opposed to the original UPDRS that measured lightheadedness as either present or absent. The MDS-UPDRS also measures nighttime sleep problems and daytime sleepiness and the yes/no sleep disturbances option is replaced from the original UPDRS. The "Complexity of Motor Fluctuations" question in the MDS-UPDRS merges the three yes/no questions related to unpredictable, predictable, and sudden OFF period from the UPDRS.

Importance of Using a Diagnosis-Specific Assessment

Parkinson's disease is generally treated as a secondary condition at Cottesmore. It is possible that patients with Parkinson's disease as a primary diagnosis are not referred to Cottesmore to receive occupational therapy services because a case has not been made as to how occupational therapists can directly or even dramatically improve the lives of persons with the disease. All occupational therapists hold the responsibility of marketing the profession to clients, payers, administrators, and other vested individuals and groups. By taking the time to administer the MDS-UPDRS to clients with Parkinson's disease, your team could gather persuasive data demonstrating improvement in Parkinson's-related symptoms as they interfere with functional activities. In doing so, you could more directly demonstrate the profession's worth and make a case for treating Parkinson's disease as a primary diagnosis at Cottesmore.

How to Obtain the MDS-UPDRS

1. Complete the Permissions Request Form: https://mds.movementdisorders.org/publications/rating_scales/request_form.php
2. Print the instructions and scoring sheet: http://www.movementdisorders.org/MDS-Files1/PDFs/MDS-UPDRS_English_FINAL.pdf

Training (Recommended)

Movement Disorder Society members: FREE
Health professional, (non-physician/non-member): $250

Movement Disorder Society Membership

$100/$175* annual fee
*Includes print subscription

Psychometric Data Overview

Psychometric Property	Definition	Data
Internal Consistency	Measures whether several items that intend to measure the same general construct yield similar scores.	Part I: Adequate (α = 0.79) Part II: Excellent (α = 0.90) Part III: Excellent (α = 0.93) Part IV: Adequate (α = 0.79)
Concurrent Validity	How closely the outcomes of one measure are similar to the outcomes of another measure.	Part I: Excellent (r = 0.76) Part II: Excellent (r = 0.92) Part III: Excellent (r = 0.96) Part IV: Excellent (r = 0.89)
Floor/Ceiling Effects	When either really low or high scores are more common between individuals because the scale doesn't distinguish differences at the extreme ends.	Part I: Adequate (lowest 0.1%/highest 0.8%) Part II: Adequate (lowest 0.1%/highest 0.7%) Part III: Adequate (lowest 0.1%/highest 0.2%) Part IV: Poor (floor effect but no ceiling effect)

Psychometric data on the MDS-UPDRS is still limited. Construct validity, content validity, test-retest reliability, minimally clinically important difference, minimal detectable change, responsiveness, and standard error of the measurement have not been established.

References

Goetz, C., Tilley, B., Shaftman, S., Stebbins, G., Fahn, S., Martin-Martinez, ... LaPelle, N. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. Movement Disorders, 23, 2129-2170.

Raad, J. (2014). Rehab measures: Movement disorder society sponsored Unified Parkinson's Disease Rating Scale revision. Retrieved from http://www.rehabmeasures.org/Lists/RehabMeasures/DispForm.aspx?ID=1134

Appendix D

Items and costs associated with creating a CAHAI administration manual

KIT ASSEMBLY FOR THE CAHAI-9

Included below are a list of the items needed to administer the 9-item version of the Chedoke Arm and Hand Activity Inventory. This list assumes that you buy every item new, but it is likely your clinic already has some of these items available for use. Because the test was developed in Canada, many of the measurements are in metric. When possible, American measurements have also been provided. A complete list of resources is available at www.cahai.ca.

Test item	**Details**	**Estimated cost**
Dycem	7.5” mat or similar	$14
200g jar of coffee	Twist-off lid, roughly 7-8oz	$6 to $12
Push button telephone		$15 to $20
12”/30cm ruler		$1
8.5” x 11” paper		$7 per ream (500 sheets)
Pencil	Will need sharpening and replacing	$2 per dozen
2.3L plastic pitcher with lid	2.25 quarts	$10
250mL plastic cup	Roughly 8oz	$10 per dozen
Washcloth		$2
Wash basin (24.5 cm. in diameter, height 8 cm.)	9.5” diameter by 3” high, roughly 4 qt capacity	$8 to $12
Pull on vest with 5 buttons (one side male & one side female)	See pattern in manual	$5 (1 yard of fabric 45” wide)
Bath towel (65cm X 100cm)		$5
75 ml toothpaste with screw lid, >50% full	Roughly 2.5 ounces, will need to be replaced periodically	$1 to $5
Toothbrush		$1
Dinner plate (Melamine or heavy plastic, 25 cm. in diameter)	Roughly 10”	$5

Medium resistance putty		$10 for 6 oz
Knife and fork	Consider Goodwill or similar	$2
Built up handles the length of the utensil handle		$15 to $20 for trimmable or pack of 4
10 gallon storage bin		$20

Total cost of assembling kit from scratch: about $152

Appendix E

Follow-up survey for clinicians regarding implementation of the CAHAI

Survey:

1. On a scale of 1-10, how informative was the in-service?

1 2 3 4 5 6 7 8 9 10
(Not informative) (Extremely informative)

Comments:

_____
_____
_____
_____

2. How useful was our research to your current practice in this setting?

1 2 3 4 5 6 7 8 9 10
(Not useful) (Extremely Useful)

Comments:

_____
_____
_____
_____

3. Do you have any further questions about either the Chedoke or Movement Disorder Society's Unified Parkinson's Disease Rating Scale?

_____
_____
_____
_____

4. How likely are you and your team to adopt use of the Chedoke (CAHAI-9)?

1 2 3 4 5 6 7 8 9 10

(Not likely) (Extremely likely)

5. How likely are you and your team to adopt use of the Movement Disorder Society's Unified Parkinson's Disease Rating Scale (MDS-UPDRS)?

1 2 3 4 5 6 7 8 9 10

(Not likely) (Extremely likely)

6. Is there anything that we could do to make implementation of either assessment in your clinic easier or more practical?

7. What potential research interests or questions do you have for future UPS students?

Permission for Scholarly Use of Thesis

To properly administer the Research Repository and preserve the contents for future use, the University of Puget Sound requires certain permissions from the author(s) or copyright owner. By accepting this license, I still retain copyright to my work. I do not give up the right to submit the work to publishers or other repositories. By accepting this license, I grant to the University of Puget Sound the non-exclusive right to reproduce, translate (as defined below), and/or distribute my submission (including the abstract) worldwide, in any format or medium for non-commercial, academic purposes only. The University of Puget Sound will clearly identify my name(s) as the author(s) or owner(s) of the submission, including a statement of my copyright, and will not make any alteration, other than as allowed by this license, to my submission. I agree that the University of Puget Sound may, without changing the content, translate the submission to any medium or format and keep more than one copy for the purposes of security, back up and preservation. I also agree that authorized readers of my work have the right to use it for non-commercial, academic purposes as defined by the "fair use" doctrine of U.S. copyright law, so long as all attributions and copyright statements are retained. If the submission contains material for which I do not hold copyright and that exceeds fair use, I represent that I have obtained the unrestricted permission of the copyright owner to grant the University of Puget Sound the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission. I further understand that, if I submit my project for publication and the publisher requires the transfer of copyright privileges, the University of Puget Sound will relinquish copyright, and remove the project from its website if required by the publisher.

Name: _____ Date: _____

_____

Julia Albright, OTS

Name: _____ Date: _____

_____

Kayla Karelsen, OTS

Name: _____ Date: _____

_____

Allison Lucas, OTS